


```
RESULT 2
ENTRY B37314 #type complete
TITLE regulatory protein lcrv - Yersinia pseudotuberculosis
ORGANISM #formal_name Yersinia pseudotuberculosis
DATE 30-Jan-1993 #sequence_revision 30-Jan-1993 #text_change 17-Mar-1999

ACCESSIONS B37314
REFERENCE A37314
#authors Bergman, T.; Hakansson, S.; Forsberg, A.; Norlander, L.;
#title J. Bacteriol. (1991) 173:1607-1616
#title Analysis of the V antigen lcrGV-yopBD operon of Yersinia
#title pseudotuberculosis: evidence for a regulatory role of lcrH
#title and lcrV.
#cross-references MUID:91154114
#accession B37314
#status Preliminary
#molecule_type DNA
#residues 1-326 #label BER
#cross-references GB:M57893; NID:q155456; PID:q155458
SUMMARY #length 326 #molecular-weight 37336 #checksum 4689

Query Match 61.8%; Score 2192; DB 2; Length 326;
Best Local Similarity 96.6%; Pred. No. 0.00e+00;
Matches 315; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

DB 1 MIRAEONPOHFIEDLEKRVFOLGHSVLELVOLVKDNIDISIKYDPRKSEVFA 60
OY 196 MIRAEONPOHFIEDLEKRVFOLGHSVLELVOLVKDNIDISIKYDPRKSEVFA 255
DB 61 NRVITDDELKKLIAYFLPEDALIKGHHYDQNLONGIKRVKEFLESSPNTOMELRAEFA 120
OY 256 NRVITDDELKKLIAYFLPEDTLKGHHYDQNLONGIKRVKEFLESSPNTOMELRAEFA 315
DB 121 VHFSLTRDIDDDILKVIYDSMNHGDAKRSKRLREELAEFLAEIKYVYQAEINKHLS 180
OY 316 VHFSLTRDIDDDILKVIYDSMNHGDAKRSKRLREELAEFLAEIKYVYQAEINKHLS 375
DB 181 GGTINIHOKSIMLPKNGVYDEDEFKASAEYKILEKMPOTTIOGGEKKIVSKNLF 240
OY 376 GGTINIHOKSIMLPKNGVYDEDEFKASAEYKILEKMPOTTIOGGEKKIVSKNLF 435
DB 241 ESEKRTGALGNLKDYSYKNKNELSHPATCSPKSRPLDLVSQKTYQLSDITSRENS 300
OY 436 GSEKRTGALGNLKNYSYKNKNELSHPATCSPKSRPLDLVSQKTYQLSDITSRENS 495
DB 301 AIEALNRFIOKIDSVQRLDDTSGK 326
OY 496 AIEALNRFIOKIDSVQRLDDTSGK 521

RESULT 3
ENTRY S13008 #type complete
TITLE capsular antigen Fl precursor - Yersinia pestis
ORGANISM #formal_name Yersinia pestis
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 13-Sep-1998
ACCESSIONS S13008; S23725
REFERENCE S13008
#authors Galysov, E.E.; Smirnov, O.Y.; Karlishhev, A.V.; Volkovoy, K.I.;
#title Denesvuk, A.I.; Nazimov, I.V.; Rubtsov, K.S.; Abramov,
#title V.M.; Dalvadganz, S.M.; Zav'yalov, V.P.
#title FEBS Lett. (1990) 277:230-232
#title Nucleotide sequence of the Yersinia pestis gene encoding Fl
#title antigen and the primary structure of the protein. Putative
#title T and B cell epitopes.
#cross-references MUID:91099503
#accession S13008
#molecule_type DNA
#residues 1-170 #label GAL
#cross-references EMBL:X6196; NID:g48620; PID:g48621
```

```
GENETICS
#gene cafi
FEATURE 1-21 #domain signal sequence #status predicted #label SIGY
22-170 #product capsular antigen Fl #status predicted #label
SUMMARY #length 170 #molecular-weight 17666 #checksum 5455

Query Match 31.0%; Score 1100; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 1.52e-139;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MKKISSVIALALFGTIAANAAADLTASTATATLVEPRIRILTYEGAPITIMDNGIDT 60
OY 24 MKKISSVIALALFGTIAANAAADLTASTATATLVEPRIRILTYEGAPITIMDNGIDT 83
DB 61 ELLVGTFLGKTKTSTSVNFTDAAGDPMYLFPTSDGNNHOTTGVICKDSRFDIS 120
OY 84 ELLVGTFLGKTKTSTSVNFTDAAGDPMYLFPTSDGNNHOTTGVICKDSRFDIS 143
DB 121 PKVNGENLVGDVVLATSGODFFVRSISGKGLAKGYTDAVTYVSNQ 170
OY 144 PKVNGENLVGDVVLATSGODFFVRSISGKGLAKGYTDAVTYVSNQ 193

RESULT 4
ENTRY I58315 #type complete
TITLE WT1 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Aug-1997
ACCESSIONS I58315
REFERENCE I58315
#authors Hamilton, T.B.; Barilla, K.C.; Romanuk, P.J.
#title Nucleic Acids Res. (1995) 23:277-284
#title High affinity binding sites for the Wilms' tumour suppressor
#title protein WT1.
#cross-references MUID:95166649
#accession I58315
#status Preliminary; translated from GB/EMBL/DBJ
#molecule_type mRNA
#residues 1-168 #label RES
#cross-references GB:S75264; NID:g896246; PID:g896247
GENETICS
#gene GDB:WT1
#cross-references GDB:120496; OMIM:194070
#map_position 11p13-11p13
SUMMARY #length 168 #molecular-weight 20165 #checksum 6457

Query Match 4.7%; Score 168; DB 2; Length 168;
Best Local Similarity 66.7%; Pred. No. 9.77e-06;
Matches 20; Conservative 7; Mismatches 1; Indels 2; Gaps 1;

DB 1 MGNHHHHHHHSGHICG--RHMRVPGV 28
OY 1 MGNHHHHHHHSGHIDDDKHKMKISSV 30

RESULT 5
ENTRY S57382 #type complete
TITLE hypothetical protein YOL087c - yeast (Saccharomyces
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 12-Dec-1997
ACCESSIONS S57382; S56781; S50418
REFERENCE S57374
#authors Zumbstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer,
#title M.
#title Yeast (1995) 11:975-986
#title A 29.425 kb segment on the left arm of yeast chromosome XV
#title contains more than twice as many unknown as known open
```

[illegible]

QY 93 GGYKTTSTSVNFTDAAGDPMTLTFTSDGNNHOTTAKVIGKDSRDPDISPKVNGENLV 152

Db 81 ID 82

QY 153 GD 154

RESULT 8

ENTRY S33821 #type complete

TITLE median body protein - Giardia lamblia

ORGANISM #formal_name Giardia lamblia

DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Sep-1997

ACCESSIONS S33821

REFERENCE S33821

#authors Marshall, J.; Holberton, D.V.

#journal J. Mol. Biol. (1993) 231:521-530

#title Sequence and structure of a new coiled coil protein from a microtubule bundle in Giardia.

#cross-references MUID:93287123

#accession S33821

#status Preliminary

#molecule_type mRNA

#residues 1-857 #label MAR

SUMMARY #cross-references EMBL:X64517; NID:g312670; PID:g312671 #length 857 #molecular_weight 100583 #checksum 6805

Query Match 3.5%; Score 124; DB 2; Length 857;

Best Local Similarity 20.1%; Pred. No. 4,618-01;

Matches 43; Conservative 71; Mismatches 82; Indels 18; Gaps 17;

Db 582 ERVVEERIOKEKNNLKKORITELQOQRATVRETEMSALREKANELDGYNRRQAREH 641

QY 205 QHFTED-LEKRYVEOLTGHGSSVLEELVOLVKDKNIDISIKYDPKDESEFA-NRVID- 261

Db 642 EIMMLDALSDDKLRONRWAMEL-TELREKVOLEKLOEKARQVEMLELRKAMD 700

QY 262 DIELK-KILAYF-LPEDTILKGGHYDNOLOGIKRV-K-EFLESSPMTQ-WELFA-EWA 315

Db 701 VDTLVEEKORLEMRLAETKIVNNYDQADKARLOEQLEKMSDKLIFEMIMD-NRRL 759

QY 316 VMHFSLTADRIDDDI--LKVIVDSNMHHGDARSKREBELAELTALKTIVYQAEINKHL 373

Db 760 KLOVK-EIDDKTAN-MEK-LYEEYKLEDDQKAT 790

QY 374 SSSGTINIDKSKINIMDKNLVG-YTD-EEIFKAS 405

RESULT 9

ENTRY S15999 #type complete

TITLE fatty-acyl-CoA synthase (EC 2.3.1.86) beta chain - yeast

ORGANISM #formal_name Yarrowia lipolytica

DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998

ACCESSIONS S15999

REFERENCE S15999

#authors Koeltig, H.; Rotner, G.; Beck, K.F.; Schweizer, M.; Schweizer, E.

#journal Mol. Gen. Genet. (1991) 226:310-314

#title The pentafunctional FAS1 genes of Saccharomyces cerevisiae and Yarrowia lipolytica are co-linear and considerably longer than previously estimated.

#cross-references MUID:91238709

#accession S15999

#molecule_type DNA

#residues 1-2076 #label MOL

SUMMARY #cross-references EMBL:X59690; NID:g297854; PID:g297855

GENETICS FAS1

#gene #superfamily yeast fatty-acyl-CoA synthase beta chain

KEYWORDS acyltransferase

SUMMARY #length 2076 #molecular_weight 230223 #checksum 5483

Query Match 3.5%; Score 123; DB 2; Length 2076;

Best Local Similarity 31.3%; Pred. No. 5,778-01;

Matches 25; Conservative 22; Mismatches 28; Indels 5; Gaps 5;

Db 1268 TDRNTRIKFEFYKLMFGODSK-FEIDTDIT-ELIGSDVYISGKALDFVHANGKEAF 1325

QY 119 TSODGNHGFOTTKV-IKDSRDFDISPKVNGENLVAGDVLATGSDQDFVFRSISKSGKL 177

Db 1326 V-GRSTAGFY-FAPMDFAI 1343

QY 178 AAGKTYDAVTVYSNOEFMI 197

RESULT 10

ENTRY S48450 #type complete

TITLE hypothetical protein YIL007c - yeast (Saccharomyces cerevisiae)

ORGANISM #formal_name Saccharomyces cerevisiae

DATE 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 12-Dec-1997

ACCESSIONS S48450

REFERENCE S48442

#authors Rowley, N.

#submission submitted to the EMBL Data Library, August 1994

#accession S48450

#molecule_type DNA

#residues 1-220 #label ROW

#cross-references GB:Z47047; EMBL:Z38113; NID:g603997; PID:g763339; MIPS:YIL007c

GENETICS

#map_position 9L

SUMMARY #length 220 #molecular_weight 24846 #checksum 3126

Query Match 3.4%; Score 120; DB 2; Length 220;

Best Local Similarity 22.8%; Pred. No. 1,128+00;

Matches 34; Conservative 39; Mismatches 69; Indels 7; Gaps 7;

Db 38 KTDIETOLEAV-FSVLEQOG-IGMDSALVTPDGY-PRSDVDVQLQVTMIRKNVNLKNDLN 94

QY 213 KVRVE-QLTGHGSSVLEELVOLVKDKNIDISIKYDPKDESEFANRYITDDIELKKILA 271

Db 95 HILORSHVLLNQHFDDNMVKSNDARRN-NDQAIQYTIPTATSEVVPSPSPDKADIKV 153

QY 272 YFLPEDYILKGGHYDNOLOGIKRVKEFLESSPMTQWELR-AFMAVHFSLTADRIDDDI 330

Db 154 DDKLISTGNVHAHNSKLNQWVVMKNE 182

QY 331 LKVIIVDSNMHHGDARSKLRE-ELAEITAE 358

RESULT 11

ENTRY H64591 #type complete

TITLE signal peptidase I Helicobacter pylori (strain 26595)

ORGANISM #formal_name Helicobacter pylori

DATE 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 10-Oct-1997

ACCESSIONS H64591

REFERENCE A64520

#authors Tomb, J.F.; White, O.; Kervilave, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klein, H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush, J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.; Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Utterback, T.R.; Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.; Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karp, P.D.; Smith, H.O.; Fraser, C.M.; Venter, J.C.

#journal Nature (1997) 388:539-547

#title The complete genome sequence of the gastric pathogen Helicobacter pylori.

[illegible]

DB	174	HHHHHHHHHOPHQALEGELLHSLSGALGAMAGPDTGVSTPAHAPMATMPHQ	233
QY	3	HHHHHHHHSSGHIDDDKHKMKRISVIAI-ALFGITATANAADLTASTATATVEPA	61
DB	234	ALSMANHGLP-SHMGCMSDVDAD	256
QY	62	RITLYKEGAPITMD-NGNIDTE	84
RESULT	14		
ENTRY		T01384	#type complete
TITLE			hypothetical protein T419.3 - Arabidopsis thaliana
ORGANISM			formal_name Arabidopsis thaliana #common_name mouse-ear cress
DATE		12-Feb-1999	#sequence_revision 12-Feb-1999 #text_change 24-Mar-1999
ACCESSIONS		T01384	
REFERENCE		Z14314	
authors			Parnell, L.D.; Gnoj, L.; de la Bastide, M.; Hameed, A.; Habermann, K.; Schutz, K.; Huang, E.; Gottesman, T.; Dedhia, N.N.; McCombie, W.R.
#submission			submitted to the EMBL Data Library, May 1998
#description			Genomic sequence of BAC T419 from Arabidopsis thaliana, chromosome IV, near 16.6 cm.
#accession		T01384	
##status			translated from GB/EMBL/DBJ
##molecule_type		DNA	
##residues		1-925	##label PAR
##cross-references		EMBL:AF069442; NID:g3342970; PID:g3924595	
##experimental_source		cultivar Columbia	
GENETICS			
#map_position	4	10/3; 54/3; 96/3; 148/3; 176/3; 219/3; 292/3; 336/3; 367/3; 418/3; 515/1; 561/1; 584/2; 604/3; 761/3	
#introns		T419.3	
#note			
SUMMARY		#length 925	#molecular-weight 102857 #checksum 8428
Query Match		3.4%;	Score 120; DB 2; Length 925;
Best Local Similarity	17.4%;	Pred. No. 1.12e+00;	
Matches	16;	Conservative	43; Mismatches 29; Indels 4; Gaps 4;
565	RHAASLIIPSSINLEDDFLLRVFCDOQLPHSEFEESQVODIGRGNGNLGKRLKELINLN	624	

```
OY 181 KYTDA-VTVTSNOEFMT-RATEQNPQHFTE-DLEKRVDELTLGHGSSVLEELVQLVKDK 237
Db 625 NEASEDCDVREVG-VMTKQGVNEEDIVERL 655
OY 238 NIDISIKYDPRKDESEVFANRVITDDIELLKKI 269
```

RESULT 15

```
ENTRY A56923 #type complete
TITLE transcription factor shn - fruit fly (Drosophila melanogaster)
```

ALTERNATE_NAMES

ORGANISM

DATE

```
#formal_name Drosophila melanogaster
11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change
17-Mar-1999
```

ACCESSIONS

```
A56923
```

REFERENCE

#authors

#journal

```
Cell (1995) 81:791-800
schurril is required for Drosophila DPP signaling and encodes
a zinc finger protein similar to the mammalian
transcription factor PRDII-BFL.
```

```
#cross-references MUID:95292346
```

```
#accession A56923
```

```
#status
```

```
preliminary; nucleic acid sequence not shown; not
compared with conceptual translation
```

```
##molecule_type mRNA
```

```
##residues 1-2529 ##label GRI
```

```
##cross-references GB:I42311
```

GENETICS

```
gene FlyBase:shn
```

```
##cross-references FlyBase:FBgn0003396
```

```
KEYWORDS DNA binding; transcription factor; zinc finger
```

```
SUMMARY #length 2529 #molecular-weight 271637 #checksum 9689
```

```
Query Match 3.4%; Score 119; DB 2; Length 2529;
```

```
Best Local Similarity 24.6%; Pred. No. 1.40e+00;
```

```
Matches 32; Conservative 38; Mismatches 55; Indels 5; Gaps 5;
```

```
Db 30 ATAAATAAAATTTTAAATPTTKR-TYRETAATVTORSTNKANIAAIAAATE 88
OY 38 TIAINAAADLTASTATATLVEPARITLTKEGAPITIMDNGNIDTELLVGTLLGGYKT 97
Db 89 ATASASATATATDAT-LT-ASKAAATAAATTDAAGSNSSSKPSTRDKL-GE-VP 144
OY 98 GTTSTSVFTDAGDPMILFTSQDGNHQTFTTKVIGKDSRDPDISPKVNGENLVGDVY 157
Db 145 LPTVDSNHII 154
OY 158 LATGSQDPFV 167
```

Search completed: Sat Nov 27 15:34:00 1999
Job time : 62 secs.

|||||
QY 256 NRVITDDIELLKKILATFLPEDTILKGGHYDNOQNGIKRKEFLLESSPNTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKIVYDSNMHNGDARSKLREELAEELAEKIVSVIOAEINKHLSS 180
QY 316 VMHFSLTADRIDDDILKIVYDSNMHNGDARSKLREELAEELAEKIVSVIOAEINKHLSS 375
Db 181 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIVSINKDFL 240
QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIVSINKDFL 435
Db 241 GSEKRRGALGNLKNSSYNKDNNELSHFATTCSDKSRPLNDLVSOQTTQSLDITSRPN 300
QY 436 GSEKRRGALGNLKNSSYNKDNNELSHFATTCSDKSRPLNDLVSOQTTQSLDITSRPN 495
Db 301 AIEALNRFIOKYDSVMORLDDTSGK 326
QY 496 AIEALNRFIOKYDSVMORLDDTSGK 521

RESULT 2
ID LCRV_YERPS STANDARD: PRT: 326 AA.
AC P23994;
DT 01-MAR-1992 (REL. 21, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE VIRULENCE-ASSOCIATED V ANTIGEN (LOW CALCIUM RESPONSE LOCUS PROTEIN DE V).
GN LCRV.
OS YERSINIA PSEUDOTUBERCULOSIS.
OC PLASMID PIB1.
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-YPIII;
RX MEDLINE: 91154114.
RA BERGMAN T., HAKANSSON S., FORSBERG A., NORLANDER L., MACCELLARO A.,
RA BAECKMAN A., BOELIN I., WOLF-WATZ H.;
RT "Analysis of the v antigen lcrGVH-yopBD operon of Yersinia
RT pseudotuberculosis: evidence for a regulatory role of lcrH and
RT lcrV";
RT J. BACTERIOL. 173:1607-1616(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA ROGGENKAMP A., HESEMMANN J.;
RL SUBMITTED (JUN-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: INVOLVED IN CA(2+) REGULATION OF YOP EXPRESSION, WHICH
CC INCLUDES THE EXPORT PROCESS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M57893; G155458; -;
DR EMBL: X96802; E239899; -;
DR PIR: B37314; B37314;
KW PLASMID: ANTIGEN: VIRULENCE.
SQ SEQUENCE 326 AA; 37336 MW; 7D84E243 CRC32;

Query Match 61.8%; Score 2192; DB 1; Length 326;
Best Local Similarity 96.6%; Pred.No. 0.00e+00;
Matches 315; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Db 1 MIRAEQNPOHFIEDLEKVRVQOLTGSSVLELVOLVKDKNIDISIKYPRKDSVEYFA 60
QY 196 MIRAEQNPOHFIEDLEKVRVQOLTGSSVLELVOLVKDKNIDISIKYPRKDSVEYFA 255

Db 61 NRVITDDIELLKKILATFLPEDALILKGGHYDNOQNGIKRKEFLLESSPNTQWELRAFMA 120
QY 256 NRVITDDIELLKKILATFLPEDTILKGGHYDNOQNGIKRKEFLLESSPNTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKIVYDSNMHNGDARSKLREELAEELAEKIVSVIOAEINKHLSS 180
QY 316 VMHFSLTADRIDDDILKIVYDSNMHNGDARSKLREELAEELAEKIVSVIOAEINKHLSS 375
Db 181 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOEETEKIVSINKDFL 240
QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSSEKKIVSINKDFL 435
Db 241 ESEKRRGALGNLKNDSYNKDNNELSHFATTCSDKSRPLNDLVSOQTTQSLDITSRPN 300
QY 436 GSEKRRGALGNLKNSSYNKDNNELSHFATTCSDKSRPLNDLVSOQTTQSLDITSRPN 495
Db 301 AIEALNRFIOKYDSVMORLDDTSGK 326
QY 496 AIEALNRFIOKYDSVMORLDDTSGK 521

RESULT 3
ID CAPL_YERPE STANDARD: PRT: 170 AA.
AC P26948;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE F1 CAPSULE ANTIGEN PRECURSOR.
GN CAPL.
OS YERSINIA PESTIS.
OC PLASMID PFRA.
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 91099503.
RA GALIYOV E.E., SMIRNOV O.Y., KARLISHEV A.V., VOLKOVY K.I.,
RA DENESTYUK A.I., NAZIMOV I.V., ROBERTOV K.S., ABRAMOV V.M.,
RA DALADYANZ S.M., ZAV'YALOV V.P.;
RT "Nucleotide sequence of the yersinia pestis gene encoding F1 antigen
RT and the primary structure of the protein. Putative T and B cell
RT epitopes";
RT FEBS LETT. 277:230-232(1990).
RN [2]
RP FBS LEFT. 277:230-232(1990).
CC -1- SUBCELLULAR LOCATION: CAPSULE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X61996; G48621; -;
DR PIR: S13008; S13008.
KW PLASMID: SIGNAL.
FT SIGNAL 1 21
FT CHAIN 22 170
FT DOMAIN 100 150
FT FT
FT FT
SQ SEQUENCE 170 AA; 17666 MW; 35432285 CRC32;

Query Match 31.0%; Score 1100; DB 1; Length 170;
Best Local Similarity 100.0%; Pred.No. 5.16e-166;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 MKKISSVIAIALFETIATANADITASTATATVLEPARITLYYKEGAPITIMDNGINDT 60
QY 24 MKKISSVIAIALFETIATANADITASTATATVLEPARITLYYKEGAPITIMDNGINDT 83
Db 61 ELIVGTLTIGYKGTSTSTVNFDAAGDPVYLFTSODGNNHOFTRKVGKSDRDIS 120
QY 84 ELIVGTLTIGYKGTSTSTVNFDAAGDPVYLFTSODGNNHOFTRKVGKSDRDIS 143

Db 121 PKVNGENVGDVVLATGSDFFVRSIGSGKGLAKGYDVAIVYSNQ 170
144 PKVNGENVGDVVLATGSDFFVRSIGSGKGLAKGYDVAIVYSNQ 193
OY 144 PKVNGENVGDVVLATGSDFFVRSIGSGKGLAKGYDVAIVYSNQ 193

RESULT 4 STANDARD: PRT: 168 AA.
ID FMS3_ECOLI
AC P15488;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CS3 FIMBRIAL SUBUNIT A PRECURSOR (CS3 PILIN).
OS ESCHERICHIA COLI.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC ESCHERICHIA.
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.
RC STRAIN-PB176;
RX MEDLINE; 90158116.
RA JALAJAKUWARI M.B., THOMAS C.J., HALTER R., MANNING P.A.;
RT "Genes for biosynthesis and assembly of CS3 pilin of CFA/II
enterotoxigenic Escherichia coli: novel regulation of pilus
production by bypassing an amber codon."
RT MOL. MICROBIOL. 3:1683-1693(1989).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE; 89032631.
RA BOYLAN M., SMYTH C.J., SCOTT J.R.;
RT "Nucleotide sequence of the gene encoding the major subunit of CS3
fimbriae of enterotoxigenic Escherichia coli."
RT INFECT. IMMUN. 56:3297-3300(1988).
RN [3]
RP SEQUENCE OF 23-45.
RC STRAIN-E9034A;
RX MEDLINE; 90036735.
RA HALL R.H., MANEVAL D.R. JR., COLLINS J.H., THEIBERT J.L.,
RA LEVINE M.M.;
RT "Purification and analysis of colonization factor antigen I, coli
RT surface antigen 1, and coli surface antigen 3 fimbriae from
RT enterotoxigenic Escherichia coli."
RT J. BACTERIOL. 171:6372-6374(1989).
RL -1- FUNCTION: FIMBRIAE (ALSO CALLED PILI), POLAR FILAMENTS RADIATING
FROM THE SURFACE OF THE BACTERIUM TO A LENGTH OF 0.5-1.5
MICROMETERS AND NUMBERING 100-300 PER CELL, ENABLE BACTERIA TO
COLONIZE THE EPITHELIUM OF SPECIFIC HOST ORGANS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X16944; G41161;
DR EMBL; M35657; G145627;
DR PIR; A34952; A34952;
DR PIR; S07904; S07904.
KW FIMBRIA; SIGNAL.
FT SIGNAL 1 22
FT CHAIN 23 168 CS3 FIMBRIAL SUBUNIT A.
FT CONFLICT 83 83 N -> S (IN REF. 2).
SQ SEQUENCE 168 AA; 17491 MW; 4D5CD89A CRC32;

Query Match 3.6%; Score 127; DB 1; Length 168;
Best Local Similarity 24.1%; Pred. No. 3.27e-02;
Matches 41; Conservative 47; Mismatches 72; Indels 10; Gaps 10;

Db 1 MLKIKYLLIGSLTSMSSYSALAGPILTKELANVSPALDITMAPQDNLSTNG-V 59
OY 24 MKRTISSV-IATALFGTIA-TANADLTASTATATVETPARITLYKEGADITIMDNGNI 81

Db 60 -SNTLVGLTLSTSDTIVSIASNTSDISKNGT-VTEAHEINNSAFAT-ISTDNANI 116
82 DTELLVGLTLGKTKGTTTS-TSVNFTDAAGDPMYLTFTSQDNNHOFTRKVIKSDRDF 140
OY 117 TLID-KNAGNTIV-KTNGSOLPNTNPLKFTTEGNEHLVSGNRRANTIT 164
OY 141 DISPKVNGENVGDVVLATGSDFFVRSIGSGKGLAKGYDVAIVYSNQ 193

RESULT 5 STANDARD: PRT: 489 AA.
ID ARPA YEAST
AC P60428;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE ACTIN-LIKE PROTEIN ARPA.
GN ARPA OR ACT3 OR X10081C OR J1012.
OS SACCCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCCHAROMYCETALES;
OC SACCCHAROMYCETACEAE; SACCCHAROMYCES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94336725.
RA HARARYA M., KARMAN A., WINTERSBERGER U.;
RT "An essential gene of Saccharomyces cerevisiae coding for an actin-
RT related protein."
RT PROC. NATL. ACAD. SCI. U.S.A. 91:8258-8262(1994).
RN [2]
RP ERRATUM.
RA HARARYA M., KARMAN A., WINTERSBERGER U.;
RT "PROC. NATL. ACAD. SCI. U.S.A. 91:10757-10757(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RX MEDLINE; 96093911.
RA MTOGA T., SCHAAF-GERSTENSCHLAGER I., CHALWATZIS N., BAUR A.,
RA BOLES E., FOURNIER C., SCHMITT S., VELTEN C., WILHELM N.,
RA ZIMMERMANN F.K.;
RT "Sequence analysis of a 3.1 kb fragment from the left arm of
RT Saccharomyces cerevisiae chromosome X, including putative proteins
RT with leucine zippers, a fungal Zn(II)-2-Cys6 binuclear cluster domain
RT and a putative alpha 2-SCB-alpha 2 binding site."
RL YEAST 11:681-689(1995).
CC -1- SIMILARITY: TO ACTIN. STRONG, TO OTHER ACTIN-LIKE PROTEINS
OF THE ARPA SUBFAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X75317; G436808;
DR EMBL; X83502; G929875;
DR EMBL; Z49356; G1008244;
DR PIR; S47608; S47608.
DR SGD; L0000027; ARPA.
DR PROSITE; PS01132; ACTINS_ACT_LIKE; 1.
DR PFAM; PF00022; actin; 2.
DR HSSP; P02568; IATN.
KW STRUCTURAL PROTEIN; CYTOSKELETON.
SQ SEQUENCE 489 AA; 54831 MW; C41435B1 CRC32;

Query Match 3.5%; Score 123; DB 1; Length 489;
Best Local Similarity 25.8%; Pred. No. 9.34e-02;
Matches 16; Conservative 21; Mismatches 24; Indels 1; Gaps 1;

Db 22 GSTTNTIGYSGSDPSPILPSVYGYKTADEGNKKIRSEOSIGIPRKDYELKPIIE-NGLV 80
OY 93 GGYKTGTSTSVNFTDAAGDPMYLTFTSQDNNHOFTRKVIKSDRDFDISPKVNGENV 152

AC P34229: (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 37, LAST ANNOTATION UPDATE)
 PT 15-DEC-1998 (REL. 38, SUBMIT BETA (EC 2.3.1.46)) [CONTAINS: 3-
 DE FACTO ACID SYNTHASE, SUBUNIT BETA]
 DE HYDROXYPALMITOYL-[LACYL-CARRIER-PROTEIN] REDUCTASE (EC 4.2.1.61);
 DE ENOYL-[LACYL-CARRIER-PROTEIN] REDUCTASE (NADH) (EC 1.3.1.9); LACYL-
 CARRIER-PROTEIN ACETYLTRANSFERASE (EC 2.3.1.38); LACYL-CARRIER-

QY	119	TSODGNHOGFTTKV-IGKDSRDFDISPKVNGENLVGGDDVVLATGSDDFVRSIGSKGKL	17
Db	1326	V--GRSTSGATV-FAPMDFAI	1343
QY	178	AAKRYDAVTYVTSNQEFMI	197
RESULT	8		


```
RT developmental control genes."
RL CYTOGENET. CELL GENET. 66:267-271(1994).
CC -1- FUNCTION: MAY PLAY A ROLE IN DETERMINING OR MAINTAINING THE
CC IDENTITIES OF A SMALL SUBSET OF VISUAL SYSTEM NEURONS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: BRAIN. PERIPHERAL SENSORY NERVOUS SYSTEM AND
CC RETINA. IN THE ADULT NERVOUS SYSTEM BRN-3.2 PREDOMINATES IN THE
CC OPTICAL, INTERMEDIATE, AND DEEP GRAY AREAS OF THE SUPERIOR
CC COLLICULUS, THE DORSAL COLUMN OF THE MESENCEPHALIC AND PONTINE
CC CENTRAL GRAY, AND THE LATERAL INTEREDUNCULAR NUCLEUS.
CC -1- DEVELOPMENTAL STAGE: IN THE CNS IT IS SELECTIVELY EXPRESSED IN
CC POSTMITOTIC, TERMINALLY DIFFERENTIATED NEURONS.
CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS
CC TO CLASS-4 POU.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S68377; G545069; -
DR EMBL; S69351; G546434; -
DR MGI; MGI:102524; POU4F2.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PFM; PFM0046; homeobox; 1.
DR PFM; PFM0157; pou; 1.
DR HSSP; P10037; 1A07.
KW NUCLEAR PROTEIN; DNA-BINDING; HOMEBOX.
FT DOMAIN 1 4
FT DOMAIN 43 51 POLY-MET.
FT DOMAIN 54 67 POLY-SER.
FT DOMAIN 69 77 POLY-GLY.
FT DOMAIN 76 85 POLY-SER.
FT DOMAIN 112 121 POLY-GLY.
FT DOMAIN 128 131 POU-IV BOX.
FT DOMAIN 157 160 POLY-HIS.
FT DOMAIN 173 184 POLY-SER.
FT DOMAIN 255 329 POLY-HIS.
FT DOMAIN 347 406 POU DOMAIN.
FT DNA_BIND 90 98 HOMEBOX.
FT CONFLICT 90 98 RRCAPLTPP -> MCAFYLOLO (IN REF. 2).
FT CONFLICT 404 405 RM -> XY (IN REF. 2).
SQ SEQUENCE 411 AA; 43172 MW; 99973657 CRC32;

Query Match 3.4%; Score 121; DB 1; Length 411;
Best Local Similarity 28.6%; Pred. No. 1,56e-01;
Matches 24; Conservative 22; Mismatches 35; Indels 3; Gaps 3;

DB 174 HHHHHHHHHHPPHALLGELLEHLSPGALGAMAGPGSTVSTPAHAPHMATNPMHQA 233
OY 3 HHHHHHHHHHSGHIDDCKHMKKISSVIAL-ALFGTATANAVALASTATATLVEPA 61
DB 234 ALSMANAGLP-SHMGCSMDVAD 256
OY 62 RTLTLYKEGAPLTIMD-NGNIDTE 84

RESULT 11
AC BR3B_HUMAN STANDARD: PRT; 410 AA.
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE BRAIN-SPECIFIC HOMEBOX/POU DOMAIN PROTEIN 3B (BRN-3B).
GN POU4F2 OR BRN3B.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-RETINA;
RX MEDLINE; 94000832.
RA XING M., ZHOU L.-J., PENG Y., EDDY R.L., SHOWS T.B., NATHANS J.;
RT "Brn-3b: a POU domain gene expressed in a subset of retinal ganglion
RT cells.";
RL NEURON 11:689-701(1993).
RN [2]
RP SEQUENCE OF 146-410 FROM N.A.
RC TISSUE-TESTIS;
RX MEDLINE; 93324388.
RA RING C.J.A., LATCHMAN D.S.;
RT "The human Brn-3b POU transcription factor shows only limited
RT homology to the Brn-3a/RDC-1 factor outside the conserved POU
RT domain.";
RL NUCLEIC ACIDS RES. 21:2946-2946(1993).
RN [3]
RP SEQUENCE OF 275-392 FROM N.A.
RX MEDLINE; 94052142.
RA BHARGAVA A.K., LI Z., WEISSMAN S.M.;
RT "Differential expression of four members of the POU family of
RT proteins in activated and phorbol 12-myristate 13-acetate-treated
RT Jurkat T cells.";
RL PROC. NATL. ACAD. SCI. U.S.A. 90:10260-10264(1993).
CC -1- FUNCTION: MAY PLAY A ROLE IN DETERMINING OR MAINTAINING THE
CC IDENTITIES OF A SMALL SUBSET OF VISUAL SYSTEM NEURONS.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: BRAIN. SEEMS TO BE SPECIFIC TO THE RETINA.
CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS
CC TO CLASS-4 POU.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U06233; G458391; -
DR EMBL; X71488; G312469; -
DR EMBL; L20434; G508494; -
DR MIM; 113725; -
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00035; POU_1; 1.
DR PROSITE; PS00465; POU_2; 1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
DR PFM; PFM0046; homeobox; 1.
DR PFM; PFM0157; pou; 1.
DR HSSP; P10037; 1A07.
KW NUCLEAR PROTEIN; DNA-BINDING; HOMEBOX.
FT DOMAIN 1 4
FT DOMAIN 43 51 POLY-MET.
FT DOMAIN 54 69 POLY-SER.
FT DOMAIN 71 80 POLY-GLY.
FT DOMAIN 81 84 POLY-SER.
FT DOMAIN 111 120 POLY-GLY.
FT DOMAIN 127 130 POU-IV BOX.
FT DOMAIN 156 159 POLY-HIS.
FT DOMAIN 172 183 POLY-SER.
FT DOMAIN 183 328 POLY-HIS.
FT DOMAIN 346 405 POU DOMAIN.
FT DNA_BIND 166 166 HOMEBOX.
FT CONFLICT 166 166 C -> S (IN REF. 2).
FT CONFLICT 183 183 MISSING (IN REF. 2).
SQ SEQUENCE 410 AA; 43159 MW; 48FFA52D CRC32;

Query Match 3.3%; Score 118; DB 1; Length 410;
Best Local Similarity 27.4%; Pred. No. 3,36e-01;
Matches 23; Conservative 23; Mismatches 35; Indels 3; Gaps 3;

DB 173 HHHHHHHHHHPPHALLGELLEHLSPGALGAMAGPGAVVSTPAHAPHMATNPMHQA 232
OY 3 HHHHHHHHHHSGHIDDCKHMKKISSVIAL-ALFGTATANAVALASTATATLVEPA 61
```

```

QY 3 HHHHHHHSSGHIDDDKMKKSSVIAI-ALFGTATANAADLASTATATLVEPA 61
Db 233 ALSMAHAGLP-SHMGCSVDVAD 255
QY 62 RITLYKRGAPITIMD-NGNIDTE 84

RESULT 12
ID BR11.BRARE STANDARD: PRT: 406 AA.
AC 090436: 090433:
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BRAIN-SPECIFIC HOMEOBOX/POU DOMAIN PROTEIN 1.1 (BRN-1.1) (FRAGMENT).
GN BRN-1.1.
OS BRACHYDANTO PERIO (ZEBRAFISH) (ZEBRA DANIO).
OC EUCARYOTA; METAZOA; CHORDATA; VERTEBRATA; ACTINOPTERYGII; NEOPTERYGII;
TELEOSTEI; EUTELEOSTEI; OSTARIOPHYSI; CYPRINIFORMES; CYPRINOIDEA;
NC CYPRINIDAE; RASBORINAE; DANIO.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 96193692.
RA SAMPATH K., STUART G.W.;
RT "Developmental expression of class III and IV POU domain genes in the
zebrafish."
RL BIOCHEM. BIOPHYS. RES. COMMUN. 219:565-571(1996).
CC -1- FUNCTION: MAY PLAY A ROLE IN SPECIFYING TERMINALLY DIFFERENTIATED
NEURONAL PHENOTYPES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN DEVELOPMENT AND IN THE
ADULT BRAIN.
CC -1- SIMILARITY: STRONG TO OTHER "POU" TRANSCRIPTION FACTORS. BELONGS
TO CLASS-3 POU.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U43898; G1399384; -
DR EMBL: U43656; G1323758; -
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS00071; HOMEOBOX_2; 1.
DR PROSITE: PS00035; POU_1; 1.
DR PROSITE: PS00465; POU_2; 1.
DR PFAM: PF00046; homeobox; 1.
DR PFAM: PF00157; pou; 1.
DR HSP: P14859; POU.
KW HOMEOBOX; DNA-BINDING; TRANSCRIPTION REGULATION; NUCLEAR PROTEIN.
FT NON-TER 1
FT DOMAIN 228 299 POU DOMAIN.
FT DNA_BIND 317 376 HOMEOBOX.
FT DOMAIN 81 92 POLY-ALA.
FT DOMAIN 203 213 POLY-HIS.
SQ SEQUENCE 406 AA; 43677 MW; DB96889B CRC32;

Query Match 3.2%; Score 114; DB 1; Length 406;
Best Local Similarity 56.5%; Pred. No. 9, 11e+01;
Matches 13; Conservative 4; Mismatches 5; Indels 1; Gaps 1;

```

```

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROBABLE GUANOSINE-3',5'-BIS(DIPHOSPHATE) 3'-PYROPHOSPHOHYDROLASE
DE (EC 3.1.7.2) ((PPGPP)ASE) (PENTA-PHOSPHATE GUANOSINE-3'-
DE PYROPHOSPHOHYDROLASE).
GN SPOT.
OS MYCOPLASMA PNEUMONIAE.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLIICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN-ATCC 29342 / M129;
RX MEDLINE: 97105885.
RA HIMMELREICH R., HILBERT H., PLAGENS H., PIRKL E., LI B.-C.,
RA HERMANN R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
pneumoniae."
RL NUCLEIC ACIDS RES. 24:4420-4449(1996).
CC -1- FUNCTION: IN EUBACTERIA PPGPP (GUANOSINE 3'-DIPHOSPHATE 5'-
DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT
COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO
CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYSES THE
DEGRADATION OF PPGPP INTO GDP. IT MAY ALSO BE CAPABLE OF
CATALYZING THE SYNTHESIS PPGPP (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: GUANOSINE-3',5'-BIS(DIPHOSPHATE) + H(2)O =
GUANOSINE-5'-DIPHOSPHATE + PYROPHOSPHATE.
CC -1- COFACTOR: MANGANESE (BY SIMILARITY).
CC -1- SIMILARITY: HIGH TO RELA PROTEIN WHICH CATALYSES PPGPP FROM
ATP AND GTP.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000043; G1674128; -
KW HYDROLASE; MANGANESE.
SQ SEQUENCE 733 AA; 86403 MW; 2654A136 CRC32;

Query Match 3.2%; Score 113; DB 1; Length 733;
Best Local Similarity 27.0%; Pred. No. 1, 1e+00;
Matches 33; Conservative 34; Mismatches 45; Indels 10; Gaps 9;

```

```

Db 201 MDNHHHHHHNOHAG-VNSHDSH 222
QY 1 MGNHHHHHHSSGHIDDDDKH 23

RESULT 13
ID SPOT_MYCPN STANDARD: PRT: 733 AA.
AC P75386;
DT 01-NOV-1997 (REL. 35, CREATED)

```

```

DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE PROBABLE GUANOSINE-3',5'-BIS(DIPHOSPHATE) 3'-PYROPHOSPHOHYDROLASE
DE (EC 3.1.7.2) ((PPGPP)ASE) (PENTA-PHOSPHATE GUANOSINE-3'-
DE PYROPHOSPHOHYDROLASE).
GN SPOT.
OS MYCOPLASMA PNEUMONIAE.
OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; MOLLIICUTES;
OC MYCOPLASMATACEAE; MYCOPLASMA.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 93066237.

```

RA GU 2., MOERSCHHELL R.P., SHERMAN F., GOLDFARB D.S.;
RT "NIP1, a gene required for nuclear transport in yeast.";
RL PROC. NATL. ACAD. SCI. U.S.A. 89:10355-10359(1992).
RN [2]
RP SEQUENCE OF 1-602 FROM N.A.
RC STRAIN-S288C / AB972.
RA CHUCHER C.M., LOUIS E.J., BARRELL B.G., RAJANDREAN M.A., WALSH S.V.;
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE OF 571-812 FROM N.A.
RC STRAIN-S288C / AB972.
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAN M.A., WALSH S.V.;
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: INVOLVED IN TRANSPORT OF PROTEINS TO THE NUCLEUS;
CC PROBABLY VIA ITS SERINE-RICH ACIDIC N-TERMINAL.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC; MAINLY.
CC -!- SIMILARITY: TO S.POMBE SPAC1E11.01C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: L02899; -!- NOT ANNOTATED_CDS.
DR EMBL: 254141; G984682; -!
DR EMBL: 249212; G798951; -!
DR PIR: A46417; A46417.
DR SGD: L0001252; NIP1.
KW TRANSPORT; PROTEIN TRANSPORT.
FT DOMAIN 15 147 ASP/GLU-RICH (ACIDIC).
FT 15 36 SER-RICH
FT CONFLICT 111 111 V -> D (IN REF. 1).
FT CONFLICT 583 583 Q -> H (IN REF. 1).
FT CONFLICT 641 641 K -> N (IN REF. 1).
FT CONFLICT 643 643 R -> N (IN REF. 1).
SQ SEQUENCE 812 AA; 93204 MW; 47993DCE CRC32;

Query Match 3.2%; Score 115; DB 1; Length 812;
Best Local Similarity 29.2%; Pred. No. 7.11e-01;
Matches 31; Conservative 28; Mismatches 39; Indels 8; Gaps 8;

Db 700 WALLPMETVLSLT-ERVQVESLKTFFSFKRFYSSFSVAKL-AETPDL-PENKVEVYL 756
QY 308 WEIRAMAMVHFSLTADRIDDLKIVDSMNH-HGD-ARSKLRELAETLAEKLYSVI 365
Db 757 QSVIA-ELIEIPAKLN-DEKTIPEVKEGDEITKLEAMVRLNKEYKI 800
QY 366 QAEINKHLSSSGTINHDKSINLMCK-NLYGYTDEEIFKASAEYKI 410

RESULT 15
ID HA34_BRELIC STANDARD: PRI: 173 AA.
AC Q99074;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE HAM34 PROTEIN.
GN HAM34.
OS BREMIA LACTUCAE (LETTUCE DOWNY MILDEN).
OC EUKARYOTA; STRAMENOPHYTES; COMYCETES; PERONOSPORALES; PERONOSPORALES;
OC BREMIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=REGEL; TISSUE=SPORE;
RX MEDLINE: 92033077.
RA "TDELSON H.S., MITCHEMORE R.W.;
RT "Highly abundant and stage-specific mRNAs in the obligate pathogen
RT Bremlia lactucaae.";
RL MOL. PLANT MICROBE INTERACT. 3:225-232(1990).

CC -!- FUNCTION: COULD BE A STRUCTURAL PROTEIN REQUIRED FOR THE
CC INFECTION PROCESS OF B.LACTUCAE.
CC -!- TISSUE SPECIFICITY: GERMINATING SPORES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X16984; G2488; -!
KW SPORELIATION; STRUCTURAL PROTEIN.
SQ SEQUENCE 173 AA; 16478 MW; 43864282 CRC32;

Query Match 3.1%; Score 109; DB 1; Length 173;
Best Local Similarity 24.3%; Pred. No. 3.06e+00;
Matches 28; Conservative 28; Mismatches 58; Indels 1; Gaps 1;

Db 1 MKFSQILVLAIAVAIAISADAPAPATTPDPAATTPPAATTTTTPPADAGTASTEQ 60
QY 25 KRISVIAIALFGTITATNADLTASTATATLVEPARITLLYKRGARITINDGNIDTE 84
Db 61 TTAGEPAAGATNGTTTPPADGTQTATAPLDATATESSASGEMTPV-GTDTSD 114
QY 85 ILVGTLLTGXYKTGTTSVNETDAAGDPWLTFTSODGNHNGFTTKVIKGDSD 139

Search completed: Sat Nov 27 15:35:00 1999
Job time : 43 secs.

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

WIDEOR (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Merch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Nov 27 15:35:18 1999; Maspar time 32.46 Seconds

Tabular output not generated. 876.148 Million cell updates/sec

Title: >US-08-699-716A-2
Description: (1-521) from US08699716A.pep
Perfect Score: 3546
Sequence: 1 MGNHHHHHHHSGHIDD.....RFICKYDSVMQRLDDTSGK 521

Scoring table: PAM 150
Gap 11

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sptrembl9
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_ricent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 51.021; Variance 127.664; scale 0.400

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2247	63.4	326	2	068697	V ANTIGEN.	0.00e+00
2	2159	60.9	324	2	087495	LCRV.	0.00e+00
3	2132	60.1	325	2	056880	V ANTIGEN (FRAGMENT).	0.00e+00
4	2128	60.0	324	2	056882	V ANTIGEN.	0.00e+00
5	2125	59.9	324	2	056879	V ANTIGEN.	0.00e+00
6	2124	59.9	324	2	056846	V ANTIGEN.	0.00e+00
7	2075	58.5	334	2	056853	V ANTIGEN (FRAGMENT).	0.00e+00
8	2074	58.5	334	2	056896	V ANTIGEN (FRAGMENT).	0.00e+00
9	653	18.4	294	2	030527	PCR.V.	1.64e-79
10	168	4.7	168	4	016256	WT1-WTLM5. TUMOR SUPPR	1.19e-02
11	133	3.8	1620	4	099986	YOTIAO (FRAGMENT).	1.69e-02
12	133	3.8	1642	4	014869	YOTIAO.	1.69e-02
13	128	3.6	882	4	060335	KIAA0594. PROTEIN (FRAG	6.05e-02
14	127	3.6	1116	3	099247	ORF YOLO87C FROM CHROM	7.78e-02
15	123	3.5	226	10	081447	CALCINEURIN B-LIKE PRO	2.11e-01
16	120	3.4	290	2	025300	SIGNAL PEPTIDASE I (LE	4.40e-01
17	119	3.4	697	10	P93203	MEP1 PROTEIN.	5.60e-01
18	121	3.4	2247	5	Q21480	COSMID M03A8.	3.45e-01
19	119	3.4	2529	5	Q24605	ZINC FINGER PROTEIN.	5.60e-01
20	116	3.3	373	2	Q55614	HIGH-AFFINITY BRANCHED	1.15e+00

21	116	3.3	625	10	Q99366	DNA-DIRECTED RNA POLIM	1.15e+00
22	116	3.3	1183	2	Q86064	CYTOTOXIN ASSOCIATED P	1.15e+00
23	117	3.3	1302	2	Q49547	LMP3 PROTEIN.	9.07e-01
24	116	3.3	2139	5	Q07569	MYOSIN HEAVY CHAIN.	1.15e+00
25	118	3.3	2344	3	Q74539	POTASSIUM SENSORY TRANSD	7.13e-01
26	112	3.2	187	2	Q67282	STROHOME SYNTHASE.	2.95e+00
27	112	3.2	197	2	Q52895	ROMPA (FRAGMENT).	2.95e+00
28	112	3.2	268	4	Q14710	POT. ORF I.	2.95e+00
29	114	3.2	417	1	Q30228	HYPOTHETICAL 47.2 KD P	1.85e+00
30	113	3.2	535	2	Q67462	RIBOSOMAL PROTEIN S01.	2.34e+00
31	113	3.2	583	11	Q08582	POTASSIUM G-PROTEIN.	2.34e+00
32	112	3.2	1182	5	Q24552	TF125 PROTEIN.	2.95e+00
33	114	3.2	1273	4	Q60316	KIAA0570 PROTEIN.	1.85e+00
34	112	3.2	3844	5	Q94648	AAR1 PROTEIN (FRAGMEN	2.95e+00
35	113	3.2	4564	5	Q77075	DHC7.	2.34e+00
36	109	3.1	197	2	Q57548	ROMPA (FRAGMENT).	5.89e+00
37	110	3.1	270	10	Q65578	HYPOTHETICAL 31.2 KD P	4.69e+00
38	110	3.1	416	1	Q30223	HYPOTHETICAL 47.2 KD P	3.72e+00
39	111	3.1	427	3	P78807	FUSION YEAST.	3.72e+00
40	111	3.1	584	4	Q00178	POTASSIUM G-PROTEIN.	3.72e+00
41	110	3.1	835	2	Q56889	INVASIN.	4.69e+00
42	110	3.1	1175	1	Q58914	HYPOTHETICAL PROTEIN M	4.69e+00
43	111	3.1	1338	5	Q23927	PROTEIN TYROSINE KINAS	3.72e+00
44	110	3.1	2288	5	Q23081	CODED FOR BY C. ELEGAN	4.69e+00
45	111	3.1	2527	5	Q24107	TRANSCRIPTION FACTOR.	3.72e+00

ALIGNMENTS

RESULT 1
ID 068697; PRELIMINARY; PRT; 326 AA.

AC 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE V ANTIGEN.
GN LCRV.
OS YERSINIA PESTIS.
OG PLASMID PC01.
OC BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KIM:
RA HU P., ELLIOTT J., MCCREADY P., SKOMRONSKI E., GARNES J.,
RA KOBAYASHI A., CARRANO A.V., BROOKER R., GARCIA E.,
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF053946; G2996251; -.
KW PLASMID.
SQ SEQUENCE 326 AA; 37240 MW; 443BDEDC CRC32;

Query Match 63.4%; Score 2247; DB 2; Length 326;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 325; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db	1	MIRAYRONPHFIEDEKRYVEOLTHGSSVLELYVQYKDKNIDISIKYDPKDSVFA	60
Qy	196	MIRAYRONPHFIEDEKRYVEOLTHGSSVLELYVQYKDKNIDISIKYDPKDSVFA	255
Db	61	NRVITDIELKILAYFPEDAIILKGHNDLQNGIKRVKFLSSPTOWELRAFMA	120
Qy	256	NRVITDIELKILAYFPEDAIILKGHNDLQNGIKRVKFLSSPTOWELRAFMA	315
Db	121	VHMFSLADRIDDLIKYIVDSNMHGDARSKLEELAEITAEIKYVYQAEINKHLS	180
Qy	316	VHMFSLADRIDDLIKYIVDSNMHGDARSKLEELAEITAEIKYVYQAEINKHLS	375
Db	181	SGTINIHDSINIMDKNLYGYDEELFKASAEKYLEKMPQTIQVDSKKYVSKDFL	240
Qy	376	SGTINIHDSINIMDKNLYGYDEELFKASAEKYLEKMPQTIQVDSKKYVSKDFL	435
Db	241	GSENRKTGALGNKNSYFNKDNNEISHPATCSDSRPLNDVSOQKTOLSDTSRFS	300

```
QY 436 GSEKRTGALGNLKNYSYNKNNELSHFATTCSDKSRPLNDVSOXTTQSLDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLLDDTSCK 326
QY 496 AIEALNRFIOKYDSVMORLLDDTSCK 521

RESULT 2
ID 087495 PRELIMINARY: PRT: 324 AA.
AC 087495;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
GN ICRV.
LCRV.
OS YERSINIA ENTEROCOLITICA.
OC BACTERIA; PROTOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-W22703;
RA TRIARTE M., LAMBERMONT I., KERBOURCH C., CORNELIS G.R.;
RT "Complete sequence of the yersinia enterocolitica pyv plasmid.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF080155; G3603358; -.
KW PLASMID.
SQ SEQUENCE 324 AA; 37296 MW; 7C9A652F CRC32;

Query Match 60.9%; Score 2159; DB 2; Length 324;
Best Local Similarity 95.4%; Pred. No. 0.00e+00;
Matches 308; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

Db 1 MIRAYEQNPQHFIEDLEKAVRVDLTGHGSSVLELVQVYDKKIDISIKYDPKDSVFA 60
QY 196 MIRAYEQNPQHFIEDLEKAVRVDLTGHGSSVLELVQVYDKKIDISIKYDPKDSVFA 255
Db 61 DRVYTDIELKKILAYFLPEDAILKGGHYDNLQNGIKRVEFLSSPTQWELRAFMA 120
QY 256 NRVYTDIELKKILAYFLPEDITLKGHYDNLQNGIKRVEFLSSPTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTISYQAEINRHLN 180
QY 316 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTISYQAEINRHLN 375
Db 181 SDTINIHDKSINLMDKNLYGTDEEIFKASAEYKILKMPQTTIKEGETEKKIYSINFL 240
QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILKMPQTTIQVGSSEKKIYSIKDFL 435
Db 241 ESENKRTGALGNLKNYSYNKNNELSHFATTCSDKSRPLNDVSOXTTQSLDITSRFS 300
QY 436 GSEKRTGALGNLKNYSYNKNNELSHFATTCSDKSRPLNDVSOXTTQSLDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLLDDT 323
QY 496 AIEALNRFIOKYDSVMORLLDDT 518

RESULT 3
ID 056880 PRELIMINARY: PRT: 325 AA.
AC 056880;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V ANTIGEN (FRAGMENT)
OS YERSINIA ENTEROCOLITICA.
OC BACTERIA; PROTOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-8081;
RX MEDLINE: 97162308.
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESEMANN J.;
RT "Medicine: 97162308.
RX MEDLINE: 97162308.
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESEMANN J.;
```

```
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen.";
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96798; E239984; -.
FT NON TER 325
SQ SEQUENCE 325 AA; 37260 MW; D6E88446 CRC32;

Query Match 60.1%; Score 2132; DB 2; Length 325;
Best Local Similarity 94.4%; Pred. No. 0.00e+00;
Matches 305; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

Db 1 MIRAYEQNPQHFIEDLEKAVRVDLTGHGSSVLELVQVYDKKIDISIKYDPKDSVFA 60
QY 196 MIRAYEQNPQHFIEDLEKAVRVDLTGHGSSVLELVQVYDKKIDISIKYDPKDSVFA 255
Db 61 DRVYTDIELKKILAYFLPEDAILKGGHYDNLQNGIKRVEFLSSPTQWELRAFMA 120
QY 256 NRVYTDIELKKILAYFLPEDITLKGHYDNLQNGIKRVEFLSSPTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTISYQAEINRHLN 180
QY 316 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTISYQAEINRHLN 375
Db 181 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILKMSQTTIKEGETEKKIYSIKNFL 240
QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILKMPQTTIQVGSSEKKIYSIKDFL 435
Db 241 ESENKRTGALGNLKNYSYNKNNELSHLTAACSDKSRPLNDVSOXTTQSLDITSRFS 300
QY 436 GSEKRTGALGNLKNYSYNKNNELSHFATTCSDKSRPLNDVSOXTTQSLDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLLDDT 323
QY 496 AIEALNRFIOKYDSVMORLLDDT 518

RESULT 4
ID 056882 PRELIMINARY: PRT: 324 AA.
AC 056882;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V ANTIGEN.
OS YERSINIA ENTEROCOLITICA.
OC BACTERIA; PROTOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y-96-P;
RX MEDLINE: 97162308.
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen.";
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96801; E239988; -.
SQ SEQUENCE 324 AA; 37162 MW; 56842D0A CRC32;

Query Match 60.0%; Score 2128; DB 2; Length 324;
Best Local Similarity 93.8%; Pred. No. 0.00e+00;
Matches 303; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

Db 1 MIRAYEQNPQHFIEDLEKAVRVDLTGHGSSVLELVQVYDKKIDISIKYDPKDSVFA 60
QY 196 MIRAYEQNPQHFIEDLEKAVRVDLTGHGSSVLELVQVYDKKIDISIKYDPKDSVFA 255
Db 61 DRVYTDIELKKILAYFLPEDAILKGGHYDNLQNGIKRVEFLSSPTQWELRAFMA 120
QY 256 NRVYTDIELKKILAYFLPEDITLKGHYDNLQNGIKRVEFLSSPTQWELRAFMA 315
Db 121 VMHFSLTADRIDDDILKVIYDSNMHHGDARSKLREELAEITAEIKTISYQAEINRHLN 180
```


QY 316 VMHSLTADRIDDDILKVIYDSNMHNGDARSKRLBELAELTAEIKIYVIOAEINKHLSS 375
Db 181 SDTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIKEGETEKIYVSIKKNFL 240
QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSEKKIYVSIKKNFL 435
Db 241 ESENKRTGALGNLKDYSYNKDNNELSHLGTCCSDKSRLDLVDSQKTTQSLDITSRFS 300
QY 436 GSENKRTGALGNLKDYSYNKDNNELSHFATTCSDKSRLDLVDSQKTTQSLDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLLDDT 323
QY 496 AIEALNRFIOKYDSVMORLLDDT 518

RESULT 5
ID 056879 PRELIMINARY: PRT: 324 AA.
AC 056879;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V ANTIGEN
OS YERSINIA ENTEROCOLITICA.
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97162308.
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESSEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen."
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96800; E239987; -
SQ SEQUENCE 324 AA; 37195 MW; E1CB300F CRC32;

Query Match 59.9%; Score 2125; DB 2; Length 324;
Best Local Similarity 93.8%; Pred. No. 0.00e+00;
Matches 303; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

Db 1 MIRAEQNPQHFIQDLKVRVEQLTGHGSSVLEELVQLVODKKIDISIKYDPKDSVEYFA 60
QY 196 MIRAEQNPQHFIQDLKVRVEQLTGHGSSVLEELVQLVODKKIDISIKYDPKDSVEYFA 255
Db 61 DRVITDDIELLKKIILAYFLPEDAILKGGHYDNLQNGIKRVEKEFLESSPNTQWELRAFMA 120
QY 256 NRVITDDIELLKKIILAYFLPEDAILKGGHYDNLQNGIKRVEKEFLESSPNTQWELRAFMA 315
Db 121 VMHSLTADRIDDDILKVIYDSNMHNGDARSKRLBELAELTAEIKIYVIOAEINKHLSS 180
QY 316 VMHSLTADRIDDDILKVIYDSNMHNGDARSKRLBELAELTAEIKIYVIOAEINKHLSS 375
Db 181 SDTINIHDKSINLMDKNLYGTNEEIFKASAEYKILEKMSQTTIKEGETEKIYVSIKKNFL 240
QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSEKKIYVSIKKNFL 435
Db 241 ESENKRTGALGNLKDYSYNKDNNELSHLGTCCSDKSRLDLVDSQKTTQSLDITSRFS 300
QY 436 GSENKRTGALGNLKDYSYNKDNNELSHFATTCSDKSRLDLVDSQKTTQSLDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLLDDT 323
QY 496 AIEALNRFIOKYDSVMORLLDDT 518

OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Y-108-P;
RX MEDLINE: 97162308.
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESSEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen."
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96796; E239985; -
SQ SEQUENCE 324 AA; 37194 MW; 5BD357FF CRC32;

Query Match 59.9%; Score 2124; DB 2; Length 324;
Best Local Similarity 94.1%; Pred. No. 0.00e+00;
Matches 304; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

Db 1 MIRAEQNPQHFIQDLKVRVEQLTGHGSSVLEELVQLVODKKIDISIKYDPKDSVEYFA 60
QY 196 MIRAEQNPQHFIQDLKVRVEQLTGHGSSVLEELVQLVODKKIDISIKYDPKDSVEYFA 255
Db 61 NRVITDDIELLKKIILAYFLPEDAILKGGHYDNLQNGIKRVEKEFLESSPNTQWELRAFMA 120
QY 256 NRVITDDIELLKKIILAYFLPEDAILKGGHYDNLQNGIKRVEKEFLESSPNTQWELRAFMA 315
Db 121 VMHSLTADRIDDDILKVIYDSNMHNGDARSKRLBELAELTAEIKIYVIOAEINKHLSS 180
QY 316 VMHSLTADRIDDDILKVIYDSNMHNGDARSKRLBELAELTAEIKIYVIOAEINKHLSS 375
Db 181 SDTINIHDKSINLMDKNLYGTNEEIFKASAEYKILEKMSQTTIKEGETEKIYVSIKKNFL 240
QY 376 SGTINIHDKSINLMDKNLYGTDEEIFKASAEYKILEKMPOTTIOVDSEKKIYVSIKKNFL 435
Db 241 ESENKRTGALGNLKDYSYNKDNNELSHLGTCCSDKSRLDLVDSQKTTQSLDITSRFS 300
QY 436 GSENKRTGALGNLKDYSYNKDNNELSHFATTCSDKSRLDLVDSQKTTQSLDITSRFS 495
Db 301 AIEALNRFIOKYDSVMORLLDDT 323
QY 496 AIEALNRFIOKYDSVMORLLDDT 518

RESULT 7
ID 056853 PRELIMINARY: PRT: 334 AA.
AC 056853;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE V ANTIGEN (FRAGMENT)
OS YERSINIA ENTEROCOLITICA.
OC BACTERIA: PROTEOBACTERIA; GAMMA SUBDIVISION; ENTEROBACTERIACEAE;
OC YERSINIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-WA-314;
RX MEDLINE: 97162308.
RA ROGGENKAMP A., GEIGER A.M., LEITRITZ L., KESSLER A., HESSEMANN J.;
RT "Passive immunity to infection with Yersinia spp. mediated by
RT anti-recombinant V antigen is dependent on polymorphism of V
RT antigen."
RL INFECT. IMMUN. 65:446-451(1997).
DR EMBL: X96796; E239982; -
FT NON_TER 334 334
SQ SEQUENCE 334 AA; 38241 MW; 71229360 CRC32;

Query Match 58.5%; Score 2075; DB 2; Length 334;
Best Local Similarity 93.1%; Pred. No. 0.00e+00;
Matches 309; Conservative 10; Mismatches 4; Indels 9; Gaps 3;

Db 1 MIRAEQNPQHFIQDLKVRVEQLTGHGSSVLEELVQLVODKKIDISIKYDPKDSVEYFA 60
QY 196 MIRAEQNPQHFIQDLKVRVEQLTGHGSSVLEELVQLVODKKIDISIKYDPKDSVEYFA 255


```

|||||
OY 1 MGHHHHHHHHSGHITDDDKHMKISSV 30

RESULT 11
ID 09996 PRELIMINARY: PRT: 1620 AA.
AC 09996:
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE YOTIAO (FRAGMENT).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA WU X., GRAVES T., BRADSHAW H.;
RT "The sequence of Homo sapiens BAC clone RG293F11."
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL SUBMITTED (SEP-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AC000066; G3645944;
FT NON-TER 1
FT SEQUENCE 1620 AA; 188793 MW; 2A1E553C CRC32;

Query Match 3.8%; Score 133; DB 4; Length 1620;
Best Local Similarity 21.9%; Pred. No. 1.69e-02;
Matches 57; Conservative 79; Mismatches 102; Indels 22; Gaps 20;

DB 247 ACOQILTHOQOLEEDHLEDYQKK-EDFTMQISFLOEKIKYEMEDQKVENSK-EE 304
OY 191 SNOEFMIRAYE-QNPOHFIEDLEKRVEDLTGHGSSVLEEL-V-QLVKKNIDISIKYDP 247
DB 305 IOEKETIEELNTKIIIEEKKTELKDKLTTADKLGLGELOEQIVQKNOEIKNMKLELTNS 364
OY 248 RKOSEFARVITDDIELLKILAYF--LPEDTIKGGHYDQLO-N-GIKRVK-EFLS 302
DB 365 KOKEROSSEIKOLMGTYE-ELOKRNKDSOFETDVOEMEQETQKLEQLRAELDEMG 423
OY 303 SPNT-QW-ELRAFNAVHFSL-TADRIDDLIKYIVSMNHGDR-SKLRELAELTA 357
DB 424 Q-QIVQMKOELIROHMAOMEKTRHKG-E-MENALRSYNTIV-NED-QIKLMVAINE 478
OY 358 ELKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGTDEIFKASAEKYLE-KMPQ 416
DB 479 LNIKLQDNTSOKEXKEELG 498
OY 417 TTIOVGSSEKIVSIKDFLG 436

RESULT 12
ID 01486 PRELIMINARY: PRT: 1642 AA.
AC 01486:
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE YOTIAO
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA LIN J.M., WISZYNSKI M., MADHAVAN R., SEALOCK R., KIM J.U., SHENG M.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF026245; G2623068;
FT NON-TER 1
FT SEQUENCE 1642 AA; 191023 MW; 99E0E48C CRC32;

```

```

Query Match 3.8%; Score 133; DB 4; Length 1642;
Best Local Similarity 21.9%; Pred. No. 1.69e-02;
Matches 57; Conservative 79; Mismatches 102; Indels 22; Gaps 20;

DB 278 ACOQILTHOQOLEEDHLEDYQKK-EDFTMQISFLOEKIKYEMEDQKVENSK-EE 335
OY 191 SNOEFMIRAYE-QNPOHFIEDLEKRVEDLTGHGSSVLEEL-V-QLVKKNIDISIKYDP 247
DB 336 IOEKETIEELNTKIIIEEKKTELKDKLTTADKLGLGELOEQIVQKNOEIKNMKLELTNS 395
OY 248 RKOSEFARVITDDIELLKILAYF--LPEDTIKGGHYDQLO-N-GIKRVK-EFLS 302
DB 396 KOKEROSSEIKOLMGTYE-ELOKRNKDSOFETDVOEMEQETQKLEQLRAELDEMG 454
OY 303 SPNT-QW-ELRAFNAVHFSL-TADRIDDLIKYIVSMNHGDR-SKLRELAELTA 357
DB 455 Q-QIVQMKOELIROHMAOMEKTRHKG-E-MENALRSYNTIV-NED-QIKLMVAINE 509
OY 358 ELKIYVIOAEINKHLSSSGTINIHDKSINLMDKNLYGTDEIFKASAEKYLE-KMPQ 416
DB 510 LNIKLQDNTSOKEXKEELG 529
OY 417 TTIOVGSSEKIVSIKDFLG 436

RESULT 13
ID 060335 PRELIMINARY: PRT: 882 AA.
AC 060335:
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE KIA0594 PROTEIN (FRAGMENT).
GN KIA0594
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RA TISSUE-BRAIN;
RX MEDLINE; 98290545.
RA NAGASE T., ISHIKAWA K., MIYAJIMA N., TANAKA A., KOTANI H., NOMURA N.,
OHARA O.;
RT "Prediction of the coding sequences of unidentified human genes. IX.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro."
RL DNA RES. 5:31-39(1998).
DR EMBL; AB011166; D1026450;
FT NON-TER 1
FT SEQUENCE 882 AA; 104906 MW; 84CC9E80 CRC32;

Query Match 3.6%; Score 128; DB 4; Length 882;
Best Local Similarity 25.4%; Pred. No. 6.05e-02;
Matches 45; Conservative 41; Mismatches 81; Indels 10; Gaps 10;

DB 431 EORHLEBOLKEIHRKLAVDGSLALRSTKLEHKON-ELROKKELLEKTKK-RQL 488
OY 343 DARSKRELAELTAEK-ITYS-VIOA-EINKHLSSSGTINIHDKSINLMDKNLYGYTDE 399
DB 489 EQ-KISSKLSGLKMEODTCNLEEEERKASTKIKEINQAKLVETYLKICTSLHTO 547
OY 400 ELFKASAEKILEKHPOTTIOVDSSEKIVS-INKPLGSEKNRITAGLNU-KNSYSYND 457
DB 548 KYDLIQNTTVISERKNKLESYMAA-SSQLRLTEQHFIELDENRORLLOCKEELMKR 603
OY 458 NNELS-HEATCSKSRPLNDIVSOIKTQSLDITSRFNFAIBALNRFIOKYSVQR 513

RESULT 14
ID 099247 PRELIMINARY: PRT: 1116 AA.
AC 099247:
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

```

```

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ORF YOL087C FROM CHROMOSOME XV.
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCAROMYCETALES;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FY1679;
RX MEDLINE; 96021609.
RA ZUMSTEIN E., PEARSON B.M., KALOGEROPOULOS A., SCHWEIZER M.;
RT "A 29.425 kb segment on the left arm of yeast chromosome XV contains
  more than twice as many unknown as known open reading frames.";
RL YEAST 11:975-986(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA ZUMSTEIN E., PEARSON B.M., KALOGEROPOULOS A., SCHWEIZER M.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RA MIPS;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; X83121; G600471; -.
DR EMBL; Z74829; E251885; -.
SQ SEQUENCE 1116 AA; 125381 MW; 099CF031 CRC32;

Query Match
Best Local Similarity 47.1%; Score 127; DB 3; Length 1116;
Matches 16; Conservative 7; Mismatches 11; Indels 0; Gaps 0;

Db 365 HHHHHHEHEONISTDAKVKKGIIDIALI 398
QY 3 HHHHHHHHSGHIDDDKHKMKISSVIALF 36

```

```

RESULT 15
ID 081447; PRELIMINARY; PRT; 226 AA.
AC 081447;
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CALCINEURIN B-LIKE PROTEIN 3.
GN CBL3.
OS ARABIDOPSIS THALIANA (MOUSE-EAR CRESS).
OC EUKARYOTA; VIRIDIPLANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;
OC EUPHYLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA; EUDICOTYLEDONS; ROSIDAE;
OC CAPPARALES; BRASSICACEAE; ARABIDOPSIS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA KUDLA J., XU Q., HARTER K., GRUISSM W., LUAN S.;
RT "Genes for calcineurin B-like proteins in Arabidopsis are
  differentially regulated by stress signals.";
RL SUBMITTED (JUL-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF076253; G3309086; -.
SQ SEQUENCE 226 AA; 26031 MW; F51F2707 CRC32;

```

```

Query Match
Best Local Similarity 3.5%; Score 123; DB 10; Length 226;
Matches 27; Conservative 28; Mismatches 42; Indels 5; Gaps 5;

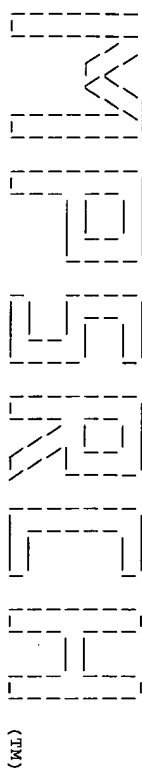
```

```

Db 125 DFFQYLKQOGFIEKQKQNVATLASGNLSDEITESTIDKTFEEADTKHGRID 184
QY 194 EFMIRAVEQNPQHFIEDLE-K-VRVEDLTGHSSVLEELVOLVQDKNIDIS-IKYPKRD 250
Db 185 KEEMRTVLNRHP-SILKNMFLQYTKDITTFPSPFVHSDVED 225
QY 251 SEVFANRVITTDIELLKKI-LATFLPEDTTLKGGHYDNLON 291

```

Search completed: Sat Nov 27 15:36:59 1999
 Job time : 101 secs.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch.p protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Nov 27 15:30:18 1999; Maspar time 20.48 Seconds

Tabular output not generated. 541.086 Million cell updates/sec

Title: >US-08-699-716A-2
Description: (1-521) from US08699716A.pep
Perfect Score: 3546
Sequence: 1 MGNHHHHHHHSSGHIDDD.....RFIOKYDSVMQRLDDTSCK 521

Scoring table: PAM 150
Gap 11

Searched: 170751 seqs, 2126608 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39

Statistics: Mean 35.735; Variance 201.794; scale 0.177

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description	Pred. No.
1	3294	92.9	W01044	Y. pestis FI/V antige	2.07e-236
2	3166	89.3	W01045	Y. pestis FI/V antige	1.17e-226
3	2242	63.2	W01040	Y. pestis V antigen	2.16e-156
4	2242	63.2	W01040	Partial lcrV (V antig	6.15e-156
5	2233	63.1	W01041	Y. pestis V antigen	1.04e-155
6	2233	63.0	R79962	Partial lcrV (V antig	3.31e-70
7	1100	31.0	W01043	Y. pestis FI antigen	3.31e-70
8	1100	31.0	R76528	Yersinia pestis cafi	3.31e-70
9	1100	31.0	W59783	Amino acid sequence o	3.31e-70
10	1100	31.0	W59782	Amino acid sequence o	1.42e-60
11	971	27.4	W59787	Nucleotide sequence o	1.42e-60
12	971	27.4	W59787	Amino acid sequence o	2.38e-60
13	968	27.3	W01042	Y. pestis FI antigen	2.38e-60
14	968	27.3	R76526	Yersinia pestis cafi	3.38e-60
15	966	27.2	R76527	Yersinia pestis cafi	7.78e-53
16	867	24.5	W59785	Amino acid sequence o	

17	738	20.8	171 34	W59786	Amino acid sequence o	2.77e-43
18	190	5.4	24 23 <td>W19779</td> <td>Polyhistidine-enterok</td> <td>3.17e-04</td>	W19779	Polyhistidine-enterok	3.17e-04
19	189	5.3	473 35 <td>W68400</td> <td>Clostridium botulinum</td> <td>3.69e-04</td>	W68400	Clostridium botulinum	3.69e-04
20	172	4.9	472 35 <td>W68393</td> <td>Clostridium botulinum</td> <td>4.61e-03</td>	W68393	Clostridium botulinum	4.61e-03
21	172	4.9	472 35 <td>W68394</td> <td>Clostridium botulinum</td> <td>4.61e-03</td>	W68394	Clostridium botulinum	4.61e-03
22	157	4.4	451 35 <td>W68398</td> <td>Clostridium botulinum</td> <td>4.11e-02</td>	W68398	Clostridium botulinum	4.11e-02
23	152	4.3	448 35 <td>W68399</td> <td>Clostridium botulinum</td> <td>8.45e-02</td>	W68399	Clostridium botulinum	8.45e-02
24	154	4.3	462 35 <td>W68397</td> <td>Clostridium botulinum</td> <td>6.34e-02</td>	W68397	Clostridium botulinum	6.34e-02
25	150	4.2	462 17 <td>R95009</td> <td>Type A neurotoxin c f</td> <td>1.12e-01</td>	R95009	Type A neurotoxin c f	1.12e-01
26	150	4.2	462 35 <td>W68380</td> <td>Clostridium botulinum</td> <td>1.12e-01</td>	W68380	Clostridium botulinum	1.12e-01
27	146	4.1	53 28 <td>W36986</td> <td>HEFT peptide.</td> <td>1.99e-01</td>	W36986	HEFT peptide.	1.99e-01
28	146	4.1	54 28 <td>W36990</td> <td>HEFT peptide.</td> <td>1.99e-01</td>	W36990	HEFT peptide.	1.99e-01
29	146	4.1	323 34	W71185	FP505 protein contain	1.99e-01
30	147	4.1	423 34	W68098	Chlamydomonas reinhar	1.73e-01
31	146	4.1	451 35 <td>W68395</td> <td>Clostridium botulinum</td> <td>1.99e-01</td>	W68395	Clostridium botulinum	1.99e-01
32	146	4.1	452 35 <td>W68396</td> <td>Clostridium botulinum</td> <td>1.99e-01</td>	W68396	Clostridium botulinum	1.99e-01
33	135	3.8	21 23 <td>W23647</td> <td>Recombinant squirrel</td> <td>9.37e-01</td>	W23647	Recombinant squirrel	9.37e-01
34	135	3.8	21 23 <td>W23650</td> <td>Recombinant squirrel</td> <td>9.37e-01</td>	W23650	Recombinant squirrel	9.37e-01
35	135	3.8	21 23 <td>W23653</td> <td>Recombinant squirrel</td> <td>9.37e-01</td>	W23653	Recombinant squirrel	9.37e-01
36	123	3.5	168 27 <td>W38342</td> <td>E. coli colonisation</td> <td>4.90e+00</td>	W38342	E. coli colonisation	4.90e+00
37	120	3.4	290 35 <td>W71474</td> <td>Helicobacter polypept</td> <td>7.36e+00</td>	W71474	Helicobacter polypept	7.36e+00
38	116	3.3	254 28 <td>W37689</td> <td>Protein sequence of c</td> <td>1.28e+01</td>	W37689	Protein sequence of c	1.28e+01
39	110	3.1	835 17 <td>R96206</td> <td>Invasin protein.</td> <td>2.79e+01</td>	R96206	Invasin protein.	2.79e+01
40	106	3.0	174 17 <td>R88503</td> <td>Borrelia burgdorferi</td> <td>4.71e+01</td>	R88503	Borrelia burgdorferi	4.71e+01
41	108	3.0	455 9	R57771	Mouse nucleobindin.	3.63e+01
42	107	3.0	561 34	W63043	Streptococcus uberis	4.14e+01
43	106	3.0	708 2	R08402	Am1105 encoded by Ida	4.71e+01
44	106	3.0	1325 23	W19540	Male-enhanced antigen	4.71e+01
45	106	3.0	2482 29 <td>W23996</td> <td>Human mitotin amino a</td> <td>4.71e+01</td>	W23996	Human mitotin amino a	4.71e+01

ALIGNMENTS

RESULT 1
ID W01044 standard; Protein; 501 AA.
AC W01044;
DT 28-DEC-1996 (first entry)
DE Y. pestis FI/V antigen fusion.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
OS Chimeric Yersinia pestis strain GB;
OS Chimeric synthetic.
FH Key location/Qualifiers
FT peptide 1..21
FT /label= Sig-peptide
FT /note= "FI antigen signal peptide"
FT protein 22..170
FT /label= FI antigen
FT /note= "mature FI antigen"
FT peptide 171..176
FT /label= linker
FT /note= "6-amino acid peptide linker"
FT protein 176..501
FT /label= V antigen
FT /note= "mature V antigen"
PN W09628551-A1.
PD 19-SEP-1996.
PR 13-MAR-1996; G00571.
PR 13-MAR-1996; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MNA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tilball RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38249.
PT Yersinia pestis V antigen and FI antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Example 3, Page 65-69; 98pp; English.
CC A fusion protein (W01044) comprises the FI antigen (see also
CC W01043) and V antigen (see also W01041) of Yersinia pestis joined
CC by a linker that allows each protein to attain its conformational
CC state. It is the product of a gene fusion (T38249) obtd. by PCR
CC amplification of Y. pestis DNA. FI/V fusion protein can be

CC expressed by gut-colonising organism transformants, to induce
CC an immune response against Y. pestis, the causative organism of
CC plague.
SQ Sequence 501 AA;

Query Match 92.9%; Score 3294; DB 19; Length 501;
Best Local Similarity 98.4%; Pred. No. 2,076-236;
Matches 493; Conservative 2; Mismatches 3; Indels 3; Gaps 2;

Db 1 mkksivaiatfgtataanaadlcastatatlvepariltlykegapitmdngnidt 60
Qy 24 MKKISSVIAIAFGTATATNADLTASTATATVTPARITLTLYEGAPITMDNGNIDT 83
Db 61 ellvgtllgykyktctstsvnfdaagqmylftsgqgnhgtctkvgkdsrdfidis 120
Qy 84 ELLVGTLLGYKYKTCTSTSVNFTDAGDPMYLTFTSQDGNHQTTKVIGKDSRDFDIS 143
Db 121 pkvgenlvgddvnlatsgqdfvrsigskgklaagkytdavtvsvngsgisegritray 180
Qy 144 PKVNGENLVGDDVNLATSGDFFVRSIGSKGKLAAGKYTDVTVSVNGE-F--MIRAY 200
Db 181 eonpqhffiedlekvrveqtlthgssvleelvglvkdknidisikydpkdssevfannrvlt 240
Qy 201 EONPQHFFIEDEKVRVEQTLTGHSVLEELVGLVKDKNIDISIKTDPKDSSEVFANRVIT 260
Db 241 ddieilkkilayflbedallkgyhynqngikrvkefleesspntqwelirafmavmhfs 300
Qy 261 DDIEILKKILAYFLBEDRTILKGGHYDNQNGIKRVKEFLESSPNTQWELRAFMAVMHFS 320
Db 301 ltadidddllkvlvdsnmhgdarskireelaetaelkylsvigaelnkhlssegctin 360
Qy 321 LTADIDDDLLKVLVDSNMHGDARSKIREELAELEKILSVIGAEINKHLSSSGCTIN 380
Db 361 ihdesinlmdknlvgtytdeeifkasaeykilekmpgttiqvgdssekkyviskdfisgenk 420
Qy 381 IHDEKINLMDKNLYGTYTEEIFKASAEYKILEKMPOTTIQVGSEKKYVISKDFLGESEK 440
Db 421 rtgalgnlksynskdneshfctcsdksrplndlvsgkctqjdsdtsrfnsaieal 480
Qy 441 RTGALGNLKSNSYNKDNELSHFATTCSDKSRPLNDLVSGKCTQJDSITSRFSNAIEAL 500
Db 481 nrftqkydswmqrllldtsqk 501
Qy 501 NRTQKYDSWMQRLLDTSQK 521

RESULT 2
ID W01045 standard; Protein: 480 AA.

AC W01045;
DT 28-DEC-1996 (first entry)
DE Y. pestis F1/V antigen fusion.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW F1 antigen; cafi.
OS Chimeric Yersinia pestis strain GB;
OS Chimeric synthetic.
FH Key Location/Qualifiers
FT protein 1..149
FT /label= F1 antigen
FT /note= "mature F1 antigen"
FT peptide 150..155
FT /label= Linker
FT /note= "6-amino acid peptide linker"
FT protein 156..480
FT /label= V antigen
FT /note= "mature V antigen"
PN W09628551-A1.
PD 19-SEP-1996.
PE 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tilball RW, Williamson ED;

DR WPI: 96-433824/43.
DR N-PSDB; T38256.
PR Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Disclosure: Page 51-55; 98pp; English.

CC A fusion protein (W01045) comprises the F1 antigen (see also
CC W01042) and V antigen (see also W01041) of Yersinia pestis joined
CC by a linker that allows each protein to attain its conformational
CC state. It is the product of a gene fusion (T38256) obtd. by PCR
CC amplification of Y. pestis DNA. F1/V fusion protein can be
CC expressed by gut-colonising organism transformants, to induce
CC an immune response against Y. pestis, the causative organism of
CC plague.

SQ Sequence 480 AA;

Query Match 89.3%; Score 3166; DB 19; Length 480;
Best Local Similarity 98.3%; Pred. No. 1,176-226;
Matches 472; Conservative 3; Mismatches 2; Indels 3; Gaps 2;

Db 1 adlstatatatlvepariltlykegapitmdngnidtelvgtllgykyktctstsv 60
Qy 45 ADLSTATATATVTPARITLTLYKEGAPITMDNGNIDTELLVGTLLGYKYKTCTSTSV 104
Db 61 nftdaagqmylftsgqgnhgtctkvgkdsrdfidispkvgenlvgddvnlatsgq 120
Qy 105 NFTDAAGDPMYLTFTSQDGNHQTTKVIGKDSRDFDISPKVNGENLVGDDVNLATGSOD 164
Db 121 fvrsgskgkklagkytdavtvsvngsgisegritrayeonpqhffiedlekvrveqtlg 180
Qy 155 FVRSGSKGKLAAGKYTDVTVSVNGE-F--MIRAYEONPQHFFIEDEKVRVEQTLG 221
Db 181 hgssvleelvglvkdknidisikydpkdssevfannrvltddieilkkilayflbedallk 240
Qy 222 HGSSVLEELVGLVKDKNIDISIKTDPKDSSEVFANRVITDDIEILKKILAYFLBEDRTILK 281
Db 241 gqhydnqngikrvkefleesspntqwelirafmavmhfsltadidddllkvlvdsnmh 300
Qy 282 GGHYNQNGIKRVKEFLESSPNTQWELRAFMAVMHFSLTADRIDDDILKVLVDSNMHN 341
Db 301 gdarskileelaetaelkylsvigaelnkhlssegctinlmdknlvgtytdeeif 360
Qy 342 GDARSKILEELAELEKILSVIGAEINKHLSSSGCTINIHDKSNLMDKNLYGTDEIF 401
Db 361 fksaeykilekmpgttiqvgdssekkyviskdfisgenkrtgalgnlksynskdnesh 420
Qy 402 FKSAEYKILEKMPOTTIQVGSEKKYISKDFLGESEKKRTGALGNLKSNSYNKDNEL 461
Db 421 shfattsdksrplndlvsgkctqjdsdtsrfnsaiealnrtfqkydswmqrllldtsqk 480
Qy 462 SHFATTCSDKSRPLNDLVSGKCTQJDSITSRFSNAIEALNRTFQKYDSWMQRLLDTSQK 521

RESULT 3
ID W01040 standard; Protein: 329 AA.

AC W01040;
DT 28-DEC-1996 (first entry)
DE Y. pestis V antigen.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW F1 antigen.
OS Yersinia pestis strain GB.
FH Key Location/Qualifiers
FT peptide 1..4
FT /note= "vector-encoded peptide"
PN W09628551-A1.
PD 19-SEP-1996.
PE 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tilball RW, Williamson ED;
DR WPI: 96-433824/43.
DR N-PSDB; T38242.

PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Disclosure: Page 25-28; 98pp: English.
CC Yersinia pestis V antigen (W01040) is capable of evoking protective
CC immune responses in animals. A combined vaccine of V and F1
CC antigens (see also W01042) can at least match the protection
CC afforded by live attenuated EV76 vaccine without any of the hazards
CC that have kept the EV vaccine from general use. The V antigen is
CC produced using the lcrV gene (see also T38242) optd. from Y.
CC pestis by PCR amplification. It can also be prepd. as a fusion with
CC F1 antigen (see also W01044-45) and expressed by attenuated AroA or
CC C Salmonella typhi as a live vaccine for long-term protection
CC against plague. Expression by gut-colonising bacterial transformants
CC produces a protective response against Y. pestis.
SQ Sequence 329 AA;
Query Match 63.2%; Score 2242; DB 19; Length 329;
Best Local Similarity 99.4%; Pred. No. 2.16e-156;
Matches 326; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
Db 3 ef-1rayeqnpqhfiedlekvrqegltghssvleelvglvkknidistkydpkdsf 61
1 |||||||
QY 194 EFMRAEQNHTEDELEKVRVQDLGHSSVLELVQLVKKNIDISIKYDPKDSF 253
1 |||||||
Db 62 faanvtiddlellkklayfipedaalkgghydnqngikrvkeflesspntqwelraf 121
1 |||||||
QY 254 FANVTITDDIELKLLAYFLPEDTILKGCHYDNQNGIKRVKEFLESSPNTQWELRAF 313
1 |||||||
Db 122 mavmhfstadriddllklyvdsmbhgdarsklreelaetaelkiysvigaeknhl 181
1 |||||||
QY 314 MAVMHFSLTADRIDDILKLYVDSMNHGDAKSKLRELELTAEIKYVIOAEINKHL 373
1 |||||||
Db 182 ssggtlnhdkslnmknlygyrdeelfkasaeykillekmpgttiqvdgsekkivskd 241
1 |||||||
QY 374 SSGGTINHDKSINLMKNLYGYRDEELFKASAEKILKEMPTTIQYDGSSEKIVSIKD 433
1 |||||||
Db 242 flgsenkrtgalgnlksysynkdnelsnfatcsdkrsrplndlvsgkttqjdsitrf 301
1 |||||||
QY 434 FLGSENRKTGALGNLKSYSYNKDNELSHFATCSDKSRPLNDLVSGKTTQJDSITRF 493
1 |||||||
Db 302 nsalealnrfqkydsvmgrrlldtsgk 329
1 |||||||
QY 494 NSALEALNRFQKYDSVMQRLDPTSGK 521
1 |||||||
RESULT 4
ID R79961 standard; Protein; 329 AA.
AC R79961;
DT 18-APR-1996 (first entry)
DE Partial lcrV (V antigen) of Y. pestis.
KW lcrV: V antigen; virulence; plague; vaccine; epitope.
OS Yersinia pestis.
PN W0952475-A1.
PD 14-SEP-1995.
PF 06-MAR-1995; G00481.
PR 08-MAR-1994; GB-004577.
PA (MINA) UK SEC FOR DEFENCE.
PI Leary SEC, Tiltall RW, Williamson ED, Leary SE;
DR WPI; 95-328268/42.
DR N-PSDB; T04222.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PT oral or parenteral vaccines for protection against plague
PS Claim 6: Page 11-13; 25pp: English.
CC R79961-62 are encoded by T04222-23 (resp.). DNA sequences (lcrV) encoding
CC all or a protective epitopic part of the mature V protein of Yersinia
CC pestis. The protein was expressed as a fusion protein with maltose
CC binding protein or glutathione-S-transferase in 3 different plasmid
CC vectors. Y. pestis is the highly virulent causative organism of plague
CC in a wide range of animals, including man. The V antigen (lcrV) is an
CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
CC The V antigen is postulated to act as a virulence antigen, and
CC transformed microorganisms contg. recombinant DNA encoding a V antigen
CC protein/peptide are useful in vaccines to protect against plague.

SQ Sequence 329 AA;
Query Match 63.2%; Score 2242; DB 15; Length 329;
Best Local Similarity 99.4%; Pred. No. 2.16e-156;
Matches 326; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
Db 3 ef-1rayeqnpqhfiedlekvrqegltghssvleelvglvkknidistkydpkdsf 61
1 |||||||
QY 194 EFMRAEQNHTEDELEKVRVQDLGHSSVLELVQLVKKNIDISIKYDPKDSF 253
1 |||||||
Db 62 faanvtiddlellkklayfipedaalkgghydnqngikrvkeflesspntqwelraf 121
1 |||||||
QY 254 FANVTITDDIELKLLAYFLPEDTILKGCHYDNQNGIKRVKEFLESSPNTQWELRAF 313
1 |||||||
Db 122 mavmhfstadriddllklyvdsmbhgdarsklreelaetaelkiysvigaeknhl 181
1 |||||||
QY 314 MAVMHFSLTADRIDDILKLYVDSMNHGDAKSKLRELELTAEIKYVIOAEINKHL 373
1 |||||||
Db 182 ssggtlnhdkslnmknlygyrdeelfkasaeykillekmpgttiqvdgsekkivskd 241
1 |||||||
QY 374 SSGGTINHDKSINLMKNLYGYRDEELFKASAEKILKEMPTTIQYDGSSEKIVSIKD 433
1 |||||||
Db 242 flgsenkrtgalgnlksysynkdnelsnfatcsdkrsrplndlvsgkttqjdsitrf 301
1 |||||||
QY 434 FLGSENRKTGALGNLKSYSYNKDNELSHFATCSDKSRPLNDLVSGKTTQJDSITRF 493
1 |||||||
Db 302 nsalealnrfqkydsvmgrrlldtsgk 329
1 |||||||
QY 494 NSALEALNRFQKYDSVMQRLDPTSGK 521
1 |||||||
RESULT 5
ID W01041 standard; Protein; 329 AA.
AC W01041;
DT 28-DEC-1996 (first entry)
DE Y. pestis V antigen.
KW Plague; vaccine; genetic immunisation; V antigen; lcrV;
KW F1 antigen.
OS Yersinia pestis strain GB.
FH Key Location/Qualifiers
FT peptide 1..4
FT /note="vector-encoded peptide"
PN W0962851-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tiltall RW, Williamson ED;
DR WPI; 96-433824/43.
DR N-PSDB; T38243.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Disclosure: Page 32-35; 98pp: English.
CC Yersinia pestis V antigen (W01041) is capable of evoking protective
CC immune responses in animals. A combined vaccine of V and F1
CC antigens (see also W01042) can at least match the protection
CC afforded by live attenuated EV76 vaccine without any of the hazards
CC that have kept the EV vaccine from general use. The V antigen is
CC produced using the lcrV gene (see also T38243) optd. from Y. pestis
CC by PCR amplification. It can also be prepd. as a fusion with F1
CC antigen (see also W01044-45) and expressed by attenuated AroA or C
CC Salmonella typhi and gut-colonising bacteria for protection against
CC plague.
SQ Sequence 329 AA;
Query Match 63.1%; Score 2236; DB 19; Length 329;
Best Local Similarity 99.7%; Pred. No. 6.15e-156;
Matches 324; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 5 1rayeqnpqhfiedlekvrqegltghssvleelvglvkknidistkydpkdsf 64
1 |||||||

QY 197 IRAYEONPOHFIEDLEKXVNEBOLTGSSVLEELVOLVYDKNDIDSIKIDPKRDSVFAN 256
DB 65 rvtlddiellkklayflpedaillkxghydnqlngikvkkefleesspntqweiafmay 124
QY 257 RVTIDDIELKKILAYFLPEDITILKGYHDNOQNGIKRVEKFELESSPNTQWELRAFMAY 316
DB 125 mhfsitaditiddillkvivdsmmhghdarsklreelaetaelklysvigaetnklhss 184
QY 317 MHFSLADRIDDDILKVIVDSNMHHGDARSKLREELAEELTAELKITYSVIOAEINKHLSSS 376
DB 185 gfinhdskeinmdknlvgvdeelfkasaeykilekmpgtivqdgsekkivsiikfij 244
QY 377 GTINIHDSKINLMCKNLVYRTDEIRKASAEYKILEKMPQTTIIOVDSSEKKIVSINDPFG 436
DB 245 senkrtgalgnlksysynkdmeishfattscksrplndivsgktqlsdtisrfnsa 304
QY 437 SENKRTGALGNLKSYSYNKDNNELSHFATTCSDKSRPLNDIVSQKTTQLSDITSRFNSA 496
DB 305 lealnrfikkydsvmgriiiddtsqk 329
QY 497 IEALNRFIOKXDSVMORLLDDTSGK 521

RESULT 6
ID R79962 standard; Protein: 329 AA.
AC R79962;
DT 18-APR-1996 (first entry)
DE Partial Lcrv (V antigen) of Y. pestis.
KW Lcrv; V antigen; virulence; plague; vaccine; epitope.
OS Yersinia pestis.
PN WO9224475-A1.
PD 14-SEP-1995.
PF 06-MAR-1995; G00481.
PR 08-MAR-1995; GB-004577.
PA (MINA) UK SEC FOR DEFENCE.
PI Leary SEC, Titball RW, Williamson ED, Leary SE;
DR WPI: 95-328266/42.
DR N-PSDB: T04223.
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in
PS oral or parenteral vaccines for protection against plague
CC Claim 6; Page 15-16; 25pp; English.
CC R79961-62 are encoded by T04222-23 (resp.), DNA sequences (lcrv encoding
CC all or a protective epitopic part of the mature V protein of Yersinia
CC pestis. The protein was expressed as a fusion protein with maltose
CC binding protein or glutathione-S-transferase in 3 different plasmid
CC vectors. Y. pestis is the highly virulent causative organism of plague
CC in a wide range of animals, including man. The V antigen (lcrv) is an
CC unstable 37.3 kDa monomeric peptide encoded on the ca. 70 kb lcr plasmid.
CC The V antigen is postulated to act as a virulence antigen, and
CC transformed microorganisms contg. recombinant DNA encoding a V antigen
CC protein/peptide are useful in vaccines to protect against plague.
SQ Sequence 329 AA;

Query Match 63.0%; Score 2233; DB 15; Length 329;
Best Local Similarity 99.4%; Pred. No. 1.04e-155;
Matches 323; Conservatve 2; Mismatches 0; Indels 0; Gaps 0;

DB 5 irayeqnpghfiedlekxvneboltgssvleelvolvdkndidlsikidpkrdsevfan 64
QY 197 IRAYEONPOHFIEDLEKXVNEBOLTGSSVLEELVOLVYDKNDIDSIKIDPKRDSVFAN 256
DB 65 rvtlddiellkklayflpedaillkxghydnqlngikvkkefleesspntqweiafmay 124
QY 257 RVTIDDIELKKILAYFLPEDITILKGYHDNOQNGIKRVEKFELESSPNTQWELRAFMAY 316
DB 125 mhfsitaditiddillkvivdsmmhghdarsklreelaetaelklysvigaetnklhss 184
QY 317 MHFSLADRIDDDILKVIVDSNMHHGDARSKLREELAEELTAELKITYSVIOAEINKHLSSS 376
DB 185 gfinhdskeinmdknlvgvdeelfkasaeykilekmpgtivqdgsekkivsiikfij 244
QY 377 GTINIHDSKINLMCKNLVYRTDEIRKASAEYKILEKMPQTTIIOVDSSEKKIVSINDPFG 436

DB 245 senkrtgalgnlksysynkdmeishfattscksrplndivsgktqlsdtisrfnsa 304
QY 437 SENKRTGALGNLKSYSYNKDNNELSHFATTCSDKSRPLNDIVSQKTTQLSDITSRFNSA 496
DB 305 lealnrfikkydsvmgriiiddtsqk 329
QY 497 IEALNRFIOKXDSVMORLLDDTSGK 521

RESULT 7
ID W01043 standard; Protein: 170 AA.
AC W01043;
DT 28-DEC-1996 (first entry)
DE Y. pestis F1 antigen (including signal peptide).
KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
OS Yersinia pestis strain GB.
FH key Location/Qualifiers
FT peptide 1.21
FT peptide /label= Sig_peptide
PN WO9628551-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 13-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.
PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 96-433824/43.
DR N-PSDB: T38248.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Example 2; Page 61-62; 98pp; English.
CC The F1 antigen (W01043), including the signal peptide, of Yersinia
CC pestis was produced from a DNA sequence (T38248) obtd. by PCR
CC amplification (see also T38257-58) of Y. pestis DNA. Expression
CC of the F1 antigen (see also W01042) by gut-colonising organisms in
CC the form of live vaccines can be used to protect an animal,
CC including humans, against plague.
SQ Sequence 170 AA;

Query Match 31.0%; Score 1100; DB 19; Length 170;
Best Local Similarity 100.0%; Pred. No. 3.31e-70;
Matches 170; Conservatve 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 mkhissvlaiaifgfiatanaadlastataclveparitltkyegapitmdnuidt 60
QY 24 MKHISSVLAIALFGITATANAADLASTATATLVEPARITLTLYEGAPITMDNGNIDT 83
DB 61 ellvgtlilgykktgtstsvnfidaagdpmylftsqdgmngfttkvigsdrdfis 120
QY 84 ELLVGTILTLGKYGKGTSTSVNFTDAAGDPMYLFTSQDGMNGHQTTRVIGKDSRDFIS 143
DB 121 pkyvgenlyvgddvlatgsqdfivrsvsgskgklydaavvtvtsng 170
QY 144 PKYVGENMLVGDDVVLATGSQDFIVRSVSGSKGLAAGRYTDAVTVTSNQ 193

RESULT 8
ID R76528 standard; Protein: 170 AA.
AC R76528;
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen.
KW Vaccine; plague; Salmonella typhimurium; Salmonella typhi;
OS Yersinia pestis.
PN WO9518231-A1.
PD 06-JUL-1995.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 95-246396/32.

DR N-PSDB: Q92819.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Disclosure: Page 20; 27pp; English.
CC The sequence represents the Y. pestis cafi (F1) antigen expressed
CC from plasmid pFORF1B. The DNA construct can be used to transform
CC human or animal gut colonizing microorganisms, specifically
CC attenuated Salmomella typhimurium or Salmomella typhi. The
CC transformed microorganisms can be used as live/attenuated vaccines
CC which induce immune responses at mucosal surfaces. The vaccines
CC provide protection against infection with Y. pestis, and are
CC parenterally and orally active vaccines offering protection
CC against bubonic and pneumonic plague.
Sequence 170 AA;
Query Match 31.0%; Score 1100; DB 14; Length 170;
Best Local Similarity 100.0%; Pred. No. 3,31e-70;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mkhissvialaifgriatanaadlaastatataleparitltkyegapitmdgnid 60
OY 24 MKKISSVIAIAIFGRIATANADLTAATATATLVEPARITLTKEGAPITMDGNIDT 83
Db 61 ellvgtlltgykvtgtsvnaftdaagdmlylttsqgdnmhfctkvigkdsrddis 120
OY 84 ELLVGTLLTGGYKVTGTSVNFDAAGDPMYLTFTSODGNHOFPTTKVIGKDSRDFDIS 143
Db 121 pkvgenlvgdvvlatsqgdfvrsiskgkilaagkytadvtyvsnq 170
OY 144 PKVNGENLVGDVVLATSQGDFVRSISKGKILAAGKYTDAVTVTVSNQ 193

RESULT 9
ID W59783 standard; Protein; 170 AA.
AC W59783;
DT 26-OCT-1998 (first entry)
DE Amino acid sequence of F1 antigen nYPf1sec510.
KW F1 antigen; plasmid; vaccine; plague; ds.
OS Yersinia pestis.
FS Key location/Qualifiers
FT CDS 17..532
FT /*tag= a
FT /product= "F1 antigen"
PN W09824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1987; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR N-PSDB: VA1596.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Disclosure: Pages 53-54; 75pp; English.
CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
Sequence 170 AA;

Query Match 31.0%; Score 1100; DB 34; Length 170;
Best Local Similarity 100.0%; Pred. No. 3,31e-70;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 mkhissvialaifgriatanaadlaastatataleparitltkyegapitmdgnid 60
OY 24 MKKISSVIAIAIFGRIATANADLTAATATATLVEPARITLTKEGAPITMDGNIDT 83
Db 61 ellvgtlltgykvtgtsvnaftdaagdmlylttsqgdnmhfctkvigkdsrddis 120
OY 84 ELLVGTLLTGGYKVTGTSVNFDAAGDPMYLTFTSODGNHOFPTTKVIGKDSRDFDIS 143

Db 121 pkvgenlvgdvvlatsqgdfvrsiskgkilaagkytadvtyvsnq 170
OY 144 PKVNGENLVGDVVLATSQGDFVRSISKGKILAAGKYTDAVTVTVSNQ 193
RESULT 10
ID W59782 standard; Protein; 170 AA.
AC W59782;
DT 26-OCT-1998 (first entry)
DE Amino acid sequence of F1 antigen pYF1sec170.
KW F1 antigen; plasmid; vaccine; plague.
OS Yersinia pestis.
FS Key location/Qualifiers
FT CDS 17..532
FT /*tag= a
FT /product= "F1 antigen"
PN W09824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1987; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR N-PSDB: VA1594.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 10; Page 52; 75pp; English.
CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
Sequence 170 AA;
Query Match 31.0%; Score 1100; DB 34; Length 170;
Best Local Similarity 100.0%; Pred. No. 3,31e-70;
Matches 170; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 mkhissvialaifgriatanaadlaastatataleparitltkyegapitmdgnid 60
OY 24 MKKISSVIAIAIFGRIATANADLTAATATATLVEPARITLTKEGAPITMDGNIDT 83
Db 61 ellvgtlltgykvtgtsvnaftdaagdmlylttsqgdnmhfctkvigkdsrddis 120
OY 84 ELLVGTLLTGGYKVTGTSVNFDAAGDPMYLTFTSODGNHOFPTTKVIGKDSRDFDIS 143
Db 121 pkvgenlvgdvvlatsqgdfvrsiskgkilaagkytadvtyvsnq 170
OY 144 PKVNGENLVGDVVLATSQGDFVRSISKGKILAAGKYTDAVTVTVSNQ 193
RESULT 11
ID W59788 standard; Protein; 149 AA.
AC W59788;
DT 26-OCT-1998 (first entry)
DE Nucleotide sequence of F1 antigen pYF1mat149.
KW F1 antigen; plasmid; vaccine; plague.
OS Yersinia pestis.
FS Key location/Qualifiers
FT CDS 17..532
FT /*tag= a
FT /product= "F1 antigen"
PN W09824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1987; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR N-PSDB: VA1609.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 10; Page 63; 75pp; English.
CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
Sequence 149 AA;

Query Match 27.4%; Score 971; DB 34; Length 149;
Best Local Similarity 100.0%; Pred. No. 1,42e-60;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 adltaataatcattvparitltykegapitlmngnidtelvgtltlggyktgtstsv 60
|||||
Qy 45 adltaataatcattvparitltykegapitlmngnidtelvgtltlggyktgtstsv 104
|||||
Db 61 nftdaagdpmylfttsqdggnhgttkvlgkdsrddispkvngenlvgddvlatgsqd 120
|||||
Qy 105 nftdaagdpmylfttsqdggnhgttkvlgkdsrddispkvngenlvgddvlatgsqd 164
|||||
Db 121 ffvrsigskgkllaagkyrdavtvtsnq 149
|||||
Qy 165 ffvrsigskgkllaagkyrdavtvtsnq 193
|||||

RESULT 12
ID W59787 standard; Protein: 150 AA.
AC W59787;
DT 26-OCT-1998 (first entry)
DE Amino acid sequence of F1 antigen pYpF1mat150.
KW F1 antigen; plasmid; vaccine; plague.
OS Yersinia pestis.
PN WO9824912-A2.
PD 11-JUN-1998.
PF 04-DEC-1997; U22617.
PR 04-DEC-1996; US-767115.
PA (HESK-) HESKA CORP.
PI Haanes EJ, Osorio JE, Thomas RE;
DR WPI: 98-333331/29.
DR N-PSDB: V41600.
PT Protection of animals against plague - using nucleic acid encoding
PT antigen from Yersinia, Pasteurella and Francisella spp.
PS Claim 10; Page 60; 75pp; English.
CC This is the amino acid sequence of a Yersinia pestis F1 antigen, used
CC in the method of the invention. Plasmid and host cells are used to
CC produce recombinant antigens, especially Yersinia pestis antigens.
CC The recombinant antigens can be used in vaccines that are capable of
CC protecting an animal from contracting plague.
SQ Sequence 150 AA;

Query Match 27.4%; Score 971; DB 34; Length 150;
Best Local Similarity 100.0%; Pred. No. 1,42e-60;
Matches 149; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 adltaataatcattvparitltykegapitlmngnidtelvgtltlggyktgtstsv 61
|||||
Qy 45 adltaataatcattvparitltykegapitlmngnidtelvgtltlggyktgtstsv 104
|||||
Db 62 nftdaagdpmylfttsqdggnhgttkvlgkdsrddispkvngenlvgddvlatgsqd 121
|||||
Qy 105 nftdaagdpmylfttsqdggnhgttkvlgkdsrddispkvngenlvgddvlatgsqd 164
|||||
Db 122 ffvrsigskgkllaagkyrdavtvtsnq 150
|||||
Qy 165 ffvrsigskgkllaagkyrdavtvtsnq 193
|||||

RESULT 13
ID W01042 standard; Protein: 151 AA.
AC W01042;
DT 28-DEC-1996 (first entry)
DE Y. pestis F1 antigen.
KW Plague; vaccine; genetic immunisation; F1 antigen; cafi;
KW V antigen
OS Yersinia pestis strain GB.
PN WO9628531-A1.
PD 19-SEP-1996.
PF 13-MAR-1996; G00571.
PR 13-MAR-1995; GB-005059.
PR 15-SEP-1995; GB-018946.
PR 05-DEC-1995; GB-024825.

PA (MINA) UK SEC FOR DEFENCE.
PI Bennett AM, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
DR WPI: 96-433824/43.
DR N-PSDB: T38244.
PT Yersinia pestis V antigen and F1 antigen or their protective
PT epitopic parts - useful in vaccine for protection against plague
PS Disclosure; Page 43-45; 98pp; English.
CC Yersinia pestis F1 antigen (W01042) is capable of evoking protective
CC immune responses in animals. A combined vaccine of F1 and V
CC antigens (see also W01040-41) can at least match the protection
CC afforded by live attenuated EV76 vaccine without any of the hazards
CC that have kept the FV vaccine from general use. The F1 antigen is
CC produced using the cafi gene (see also T38244) obtd. from Y.
CC pestis by PCR amplification. It can also be prep. as a fusion with
CC V antigen (see also W01044-45) and expressed by attenuated AroA or
CC C Salmonella typhi as a live vaccine for long-term protection
CC against plague. Expression by gut-colonising bacterial transformants
CC produces a protective response against Y. pestis.
SQ Sequence 151 AA;

Query Match 27.3%; Score 968; DB 19; Length 151;
Best Local Similarity 98.0%; Pred. No. 2,38e-60;
Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 ssadltaataatcattvparitltykegapitlmngnidtelvgtltlggyktgtstsv 60
:::|||||
Qy 43 ssadltaataatcattvparitltykegapitlmngnidtelvgtltlggyktgtstsv 102
:::|||||
Db 61 svnftdaagdpmylfttsqdggnhgttkvlgkdsrddispkvngenlvgddvlatgs 120
|||||
Qy 103 svnftdaagdpmylfttsqdggnhgttkvlgkdsrddispkvngenlvgddvlatgs 162
|||||
Db 121 qdfvrsigskgkllaagkyrdavtvtsnq 151
|||||
Qy 163 qdfvrsigskgkllaagkyrdavtvtsnq 193
|||||

RESULT 14
ID R76526 standard; Protein: 151 AA.
AC R76526;
DT 17-DEC-1995 (first entry)
DE Yersinia pestis cafi (F1) antigen.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
KW bubonic plague; pneumonic plague.
OS Yersinia pestis.
PN WO9518231-A1.
PD 06-JUL-1995.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA) UK SEC FOR DEFENCE.
PI Howells A, Leary SEC, Oyston PCF, Tibball RW, Williamson ED;
DR WPI: 95-246396/32.
DR N-PSDB: Q92817.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response.
PT against Yersinia pestis, at mucosal surfaces.
PS Disclosure; Page 16; 27pp; English.
CC The sequence represents the Y. pestis cafi (F1) antigen expressed
CC from plasmid pFGal2a. The DNA construct can be used to transform
CC human or animal gut colonizing microorganisms, specifically
CC attenuated Salmonella typhimurium or Salmonella typhi. The
CC transformed microorganisms can be used as live/attenuated vaccines
CC which induce immune responses at mucosal surfaces. The vaccines
CC provide protection against infection with Y. pestis, and are
CC parenterally and orally active vaccines offering protection
CC against bubonic and pneumonic plague.
SQ Sequence 151 AA;

Query Match 27.3%; Score 968; DB 14; Length 151;
Best Local Similarity 98.0%; Pred. No. 2,38e-60;
Matches 148; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 1 ssadltaataatcattvparitltykegapitlmngnidtelvgtltlggyktgtstsv 60

WORLDWIDE
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msrch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Sat Nov 27 15:37:17 1999; MasPar time 8.24 Seconds

Tabular output not generated. 753.590 Million cell updates/sec

Title: >US-08-699-716A-2
Description: (1521) from US08699716A.pep
Perfect Score: 3546
Sequence: 1 MGHNNHHHHSSGHIDDD.....RFIOKDYVMQRLDDTSCK 521

Scoring table: PAM 150
Gap 11

Searched: 122461 segs, 11912985 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:PCIT9_COMB 4:backfiles1

Statistics: Mean 33.116; Variance 187.062; scale 0.177

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	150	4.2	462	2	US-08-405- Sequence 26, Applicati	3.84e-02
2	150	4.2	462	1	US-08-480- Sequence 26, Applicati	3.84e-02
3	144	4.1	21	2	US-08-651- Sequence 21, Applicati	9.19e-02
4	146	4.1	23	2	US-08-405- Sequence 24, Applicati	6.87e-02
5	146	4.1	23	1	US-08-480- Sequence 24, Applicati	6.87e-02
6	146	4.1	53	2	US-08-651- Sequence 19, Applicati	6.87e-02
7	146	4.1	54	2	US-08-651- Sequence 23, Applicati	6.87e-02
8	146	4.1	323	2	US-08-591- Sequence 5, Applicatio	6.87e-02
9	123	3.5	168	1	US-08-460- Sequence 5, Applicatio	1.82e+00
10	106	3.0	174	3	PCT-US95-0 Sequence 2, Applicatio	1.87e+01
11	106	3.0	174	2	US-08-719- Sequence 2, Applicatio	1.87e+01
12	106	3.0	174	1	US-08-261- Sequence 2, Applicatio	1.87e+01
13	106	3.0	2482	1	US-08-328- Sequence 6, Applicatio	1.87e+01
14	105	3.0	3248	3	PCT-US95-1 Sequence 1, Applicatio	2.14e+01
15	105	3.0	3248	1	US-08-353- Sequence 1, Applicatio	2.14e+01
16	103	2.9	257	1	US-07-781- Sequence 2, Applicatio	5.39e+01
17	98	2.8	17	2	US-08-600- Sequence 13, Applicati	5.39e+01
18	98	2.8	380	1	US-08-670- Sequence 4, Applicatio	5.39e+01
19	98	2.8	380	2	US-08-585- Sequence 4, Applicatio	5.39e+01
20	98	2.8	380	2	US-08-977- Sequence 4, Applicatio	5.39e+01
21	98	2.8	381	2	US-08-670- Sequence 4, Applicatio	5.39e+01
22	98	2.8	381	2	US-08-977- Sequence 2, Applicatio	5.39e+01
23	98	2.8	381	2	US-08-786- Sequence 3, Applicatio	5.39e+01

24	98	2.8	381	1	US-08-585- Sequence 2, Applicatio	5.39e+01
25	98	2.8	390	2	US-08-786- Sequence 1, Applicatio	5.39e+01
26	100	2.8	793	1	US-08-015- Sequence 3, Applicatio	4.14e+01
27	100	2.8	802	1	US-08-015- Sequence 1, Applicatio	4.14e+01
28	99	2.8	3567	2	US-07-642- Sequence 4, Applicatio	4.73e+01
29	96	2.7	180	1	US-08-328- Sequence 8, Applicatio	6.99e+01
30	96	2.7	183	1	US-08-157- Sequence 6, Applicatio	6.99e+01
31	96	2.7	183	3	PCT-US95-1 Sequence 7, Applicatio	6.99e+01
32	96	2.7	183	3	PCT-US95-0 Sequence 20, Applicati	6.99e+01
33	97	2.7	302	2	US-08-203- Sequence 4, Applicatio	6.14e+01
34	97	2.7	302	3	PCT-US95-0 Sequence 4, Applicatio	6.14e+01
35	97	2.7	303	3	PCT-US95-0 Sequence 2, Applicatio	6.14e+01
36	97	2.7	303	2	US-08-203- Sequence 2, Applicatio	6.14e+01
37	96	2.7	403	2	US-08-961- Sequence 2, Applicatio	6.99e+01
38	97	2.7	414	3	PCT-US92-0 Sequence 2, Applicatio	6.14e+01
39	95	2.7	544	2	US-08-694- Sequence 10, Applicati	7.96e+01
40	95	2.7	699	2	US-08-694- Sequence 16, Applicati	7.96e+01
41	97	2.7	763	2	US-08-677- Sequence 2, Applicatio	6.14e+01
42	95	2.7	926	1	US-08-387- Sequence 6, Applicatio	7.96e+01
43	95	2.7	977	2	US-08-694- Sequence 8, Applicatio	7.96e+01
44	95	2.7	1069	1	US-07-777- Sequence 9, Applicatio	7.96e+01
45	96	2.7	1338	2	US-08-728- Sequence 9, Applicatio	6.99e+01

ALIGNMENTS

RESULT	1	STANDARD;	PRT:	462 AA.
ID	US-08-405-496A-26			
XX	xxxxxxx			
AC				
DT				
XX				
DE	Sequence 26, Application US/08405496A			
CC	Sequence 26, Application US/08405496A			
CC	Patent No. 5919665			
CC	GENERAL INFORMATION:			
CC	APPLICANT: WILLIAMS, JAMES A.			
CC	TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM			
CC	TITLE OF INVENTION: NEUROTOXIN			
CC	NUMBER OF SEQUENCES: 30			
CC	CORRESPONDENCE ADDRESS:			
CC	ADDRESSEE: MEDLEN & CARROLL, LLP			
CC	STREET: 220 MONTGOMERY STREET, SUITE 2200			
CC	CITY: SAN FRANCISCO			
CC	STATE: CALIFORNIA			
CC	COUNTRY: USA			
CC	ZIP: 94104			
CC	COMPUTER READABLE FORM:			
CC	MEDIUM TYPE: Floppy disk			
CC	COMPUTER: IBM PC compatible			
CC	OPERATING SYSTEM: PC-DOS/MS-DOS			
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30			
CC	CURRENT APPLICATION DATA:			
CC	APPLICATION NUMBER: US/08/405,496A			
CC	FILING DATE: 16-MAR-1995			
CC	CLASSIFICATION: 424			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 08/329,154			
CC	FILING DATE: 25-OCT-1994			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 08/161,907			
CC	FILING DATE: 02-DEC-1993			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 07/985,321			
CC	FILING DATE: 04-DEC-1992			
CC	PRIOR APPLICATION DATA:			
CC	APPLICATION NUMBER: US 07/429,791			
CC	FILING DATE: 31-OCT-1989			
CC	ATTORNEY/AGENT INFORMATION:			
CC	NAME: INGOLTA, DIANE E.			
CC	REGISTRATION NUMBER: 40,027			

CC REFERENCE/DOCKET NUMBER: OPHD-01308
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 26:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 462 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 462 AA: 53932 MW: 1087243 CN;

Query Match 4.2%; Score 150; DB 2; Length 462;
Best Local Similarity 59.4%; Pred. No. 3,84e-02;
Matches 19; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

Db 1 MGNHHHHHHHSGHIEG--RHMASMARLS 30
QY 1 MGNHHHHHHHSGHIDDDKMKKISSVIA 32

RESULT 2
ID US-08-480-604A-26 STANDARD; PRT; 462 AA.
XX xxxxxx
DE Sequence 26, Application US/08480604A
XX Patent No. 5736139
CC GENERAL INFORMATION:
CC APPLICANT: KINK, JOHN A.
CC APPLICANT: THALEY, BRUCE S.
CC APPLICANT: PADNEY, NISHA V.
CC APPLICANT: FIRCA, JOSEPH R.
CC APPLICANT: STAFORD, DOUGLAS C.
CC TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
CC TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL, LLP
CC STREET: 220 MONTGOMERY STREET, SUITE 2200
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: UNITED STATES OF AMERICA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/480,604A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/422,711
CC FILING DATE: 14-APR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/405,496
CC FILING DATE: 16-MAR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/329,154
CC FILING DATE: 25-OCT-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/161,907
CC FILING DATE: 02-DEC-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/985,321
CC FILING DATE: 04-DEC-1992
CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/429,791
CC FILING DATE: 31-OCT-1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: INGOLIA, DIANE E.
CC REGISTRATION NUMBER: 40,027
CC REFERENCE/DOCKET NUMBER: OPHD-01763
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 26:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 462 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 462 AA: 53932 MW: 1087243 CN;

Query Match 4.2%; Score 150; DB 1; Length 462;
Best Local Similarity 59.4%; Pred. No. 3,84e-02;
Matches 19; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

Db 1 MGNHHHHHHHSGHIEG--RHMASMARLS 30
QY 1 MGNHHHHHHHSGHIDDDKMKKISSVIA 32

RESULT 3
ID US-08-651-818A-21 STANDARD; PRT; 21 AA.
XX xxxxxx
DE Sequence 21, Application US/08651818A
XX Patent No. 5948889
CC GENERAL INFORMATION:
CC APPLICANT: de Boer, Piet A.J.
CC APPLICANT: Hale, Cynthia A.
CC TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR SCREENING
CC TITLE OF INVENTION: ANTIMICROBIALS
CC NUMBER OF SEQUENCES: 25
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL
CC STREET: 220 Montgomery Street, Suite 2200
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/651,818A
CC FILING DATE:
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Carroll, Peter G.
CC REGISTRATION NUMBER: 32,837
CC REFERENCE/DOCKET NUMBER: CASE-02249
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 21 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: not relevant
CC TOPOLOGY: not relevant
CC MOLECULE TYPE: peptide

SO SEQUENCE 21 AA; 2539 MW; 2077 CN;
Query Match 4.1%; Score 144; DB 2; Length 21;
Best Local Similarity 94.4%; Pred. No. 9,19e-02;
Matches 17; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 1 MGNHHHHHHHSSGHIE 18
1 MGNHHHHHHHSSGHID 18
QY 1 MGNHHHHHHHSSGHID 18
RESULT 4
ID US-08-405-496A-24 STANDARD; PRT: 23 AA.
XX AC xxxxxx
XX DT
XX DE Sequence 24, Application US/08405496A
XX CC Sequence 24, Application US/08405496A
XX CC Patent No. 5919665
CC GENERAL INFORMATION:
CC APPLICANT: WILLIAMS, JAMES A.
CC TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM
CC TITLE OF INVENTION: NEUROTOXIN
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL, LLP
CC STREET: 220 MONTGOMERY STREET, SUITE 2200
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/405,496A
CC FILING DATE: 16-MAR-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/329,154
CC FILING DATE: 25-OCT-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/161,907
CC FILING DATE: 02-DEC-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/985,321
CC FILING DATE: 04-DEC-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/429,791
CC FILING DATE: 31-OCT-1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: INGOLTA, DIANE E.
CC REGISTRATION NUMBER: 40,027
CC REFERENCE/DOCKET NUMBER: OPHD-01308
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 24:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 23 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 23 AA; 2741 MW; 2386 CN;
Query Match 4.1%; Score 146; DB 2; Length 23;
Best Local Similarity 79.2%; Pred. No. 6.87e-02;

Matches 19; Conservative 2; Mismatches 1; Indels 2; Gaps 1;
Db 1 MGNHHHHHHHSSGHIEG--RHM 22
1 MGNHHHHHHHSSGHIDDDDKM 24
QY 1 MGNHHHHHHHSSGHIDDDDKM 24
RESULT 5
ID US-08-480-604A-24 STANDARD; PRT: 23 AA.
XX AC xxxxxx
XX DT
XX DE Sequence 24, Application US/08480604A
XX CC Sequence 24, Application US/08480604A
XX CC Patent No. 5736139
CC GENERAL INFORMATION:
CC APPLICANT: KINK, JOHN A.
CC APPLICANT: THALLEY, BRUCE S.
CC APPLICANT: PADHYE, NISHA V.
CC APPLICANT: FIRCA, JOSEPH R.
CC APPLICANT: STAFFORD, DOUGLAS C.
CC TITLE OF INVENTION: VACCINE AND ANTITOXIN FOR TREATMENT AND
CC TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE
CC NUMBER OF SEQUENCES: 32
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: MEDLEN & CARROLL, LLP
CC STREET: 220 MONTGOMERY STREET, SUITE 2200
CC CITY: SAN FRANCISCO
CC STATE: CALIFORNIA
CC COUNTRY: UNITED STATES OF AMERICA
CC ZIP: 94104
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/480,604A
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/422,711
CC FILING DATE: 14-APR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/405,496
CC FILING DATE: 16-MAR-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/329,154
CC FILING DATE: 25-OCT-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/161,907
CC FILING DATE: 02-DEC-1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/985,321
CC FILING DATE: 04-DEC-1992
CC ATTORNEY/AGENT INFORMATION:
CC NAME: INGOLTA, DIANE E.
CC REGISTRATION NUMBER: 40,027
CC REFERENCE/DOCKET NUMBER: OPHD-01763
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 705-8410
CC TELEFAX: (415) 397-8338
CC INFORMATION FOR SEQ ID NO: 24:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 23 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown

[illegible]

SEQUENCE	174 AA; 20290 MW; 153533 CN;
Query Match	3.0%; Score 106; DB 2; Length 174;
Best Local Similarity	34.5%; Pred. No. 1.87e+01;
Matches	20; Conservative 12; Mismatches 24; Indels 2; Gaps 2;
Db	115 IKGPNSIFDYLIDLSDSK-IDVAEKYGE-KARENFESYKKDKITPAVKOILKOIILAD 170
QY	219 LTGSSSVLEELVQLVADKNIDISIKTIDPKRDSVFANRVITDIDELKKIILAFLE 276
RESULT	12
ID	US-08-261-825-2
XX	STANDARD;
XX	PRT;
XX	174 AA.
XX	xxxxxx
DE	Sequence 2, Application US/08261825
CC	Sequence 2, Application US/08261825
CC	Patent No. 5558993
CC	GENERAL INFORMATION:
CC	APPLICANT: Champion, Cheryl I.
CC	APPLICANT: Loyett, Michael A.
CC	APPLICANT: Haake, David A.
CC	APPLICANT: Miller, James N.
CC	APPLICANT: Blanco, David R.
CC	TITLE OF INVENTION: CLONED Borrelia burgdorferi VIRULENCE
CC	TITLE OF INVENTION: PROTEIN
CC	NUMBER OF SEQUENCES: 8
CC	CORRESPONDENCE ADDRESSES:
CC	ADDRESSEE: Spensley Horn Jubas & Lubitz
CC	STREET: 1880 Century Park East, Suite 500
CC	City: Los Angeles
CC	STATE: California
CC	COUNTRY: USA
CC	ZIP: 90067
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: Patentin Release #1.0, Version #1.25
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/261,825
CC	FILING DATE: 17-JUN-1994
CC	CLASSIFICATION: 435
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: TURKIN, LISA A.
CC	REGISTRATION NUMBER: P-38,347
CC	REFERENCE/DOCKET NUMBER: PDS316
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (619) 455-5100
CC	TELEFAX: (619) 455-5110
CC	INFORMATION FOR SEQ ID NO: 2:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 174 amino acids
CC	TYPE: amino acid
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: protein
CC	SEQUENCE
CC	174 AA; 20290 MW; 153533 CN;
CC	Query Match
CC	3.0%; Score 106; DB 1; Length 174;
CC	Best Local Similarity 34.5%; Pred. No. 1.87e+01;
CC	Matches
CC	20; Conservative 12; Mismatches 24; Indels 2; Gaps 2;
CC	Db
CC	115 IKGPNSIFDYLIDLSDSK-IDVAEKYGE-KARENFESYKKDKITPAVKOILKOIILAD 170
CC	QY
CC	219 LTGSSSVLEELVQLVADKNIDISIKTIDPKRDSVFANRVITDIDELKKIILAFLE 276
CC	RESULT
CC	12
CC	ID
CC	US-08-328-254-6
CC	STANDARD;
CC	PRT;
CC	2482 AA.

[illegible]

QY	DB	OY	RESULT	ID	PCT	US95-16216-1	STANDARD;	PRT;	3248 AA.
411	LEKPAQTTIQV-DSSEKTIYSIKRPLGSENKRTGALGKLNKSYSYNDKNELSHFATTC-	468							
1847	ELEKAQLOGLDEAKNNYIV-IQSSVNGLIQVEDGKQKLE	1886							
469	S-DKSRPLNDLVQKTTQGLSDITSRFPNSAIEALNRFIQKYD	508							
XX	Sequence 1, Application PC/TUS9516216								
XX	GENERAL INFORMATION:								
XX	APPLICANT: Yen, Timothy J.								
XX	APPLICANT: Rattner, Jerome B.								
XX	TITLE OF INVENTION: Nucleic Acid Encoding a Transiently								
XX	TITLE OF INVENTION: Expressed Kinetochores Protein, and Methods of Use								
XX	NUMBER OF SEQUENCES: 4								
XX	CORRESPONDENCE ADDRESS:								
XX	ADDRESSEE: Dann, Dorfman, Herrell and Skillman								
XX	STREET: 1601 Market Street								
XX	CITY: Philadelphia								
XX	STATE: PA								
XX	COUNTRY: USA								
XX	ZIP: 19103-2307								
XX	COMPUTER READABLE FORM:								
XX	MEDIUM TYPE: Floppy disk								
XX	COMPUTER: IBM PC compatible								
XX	OPERATING SYSTEM: PC-DOS/MS-DOS								
XX	CURRENT APPLICATION DATA:								
XX	APPLICATION NUMBER: PCT/US95/16216								
XX	FILING DATE:								
XX	CLASSIFICATION:								
XX	PRIOR APPLICATION DATA:								
XX	APPLICATION NUMBER: US 08/353,700								
XX	FILING DATE: 09-DEC-1995								
XX	ATTORNEY/AGENT INFORMATION:								
XX	NAME: Reed, Janet E.								
XX	REGISTRATION NUMBER: 36,252								
XX	TELEPHONE: (215) 563-4100								
XX	TELEFAX: (215) 563-4044								
XX	INFORMATION FOR SEQ ID NO: 1:								
XX	SEQUENCE CHARACTERISTICS:								
XX	LENGTH: 3248 amino acids								
XX	TYPE: amino acid								
XX	STRANDEDNESS: not relevant								
XX	TOPOLOGY: not relevant								
XX	MOLECULE TYPE: protein								
XX	HYPOTHETICAL: NO								
XX	ANTI-SENSE: NO								
XX	SEQUENCE 3248 AA: 372207 MW: 51689535 CN:								
XX	Query Match 3.0%; Score 105; DB 3; Length 3248;								
XX	Best Local Similarity 18.4%; Pred. NO. 2,14e+01;								
XX	Matches 69; Conservative 112; Mismatches 171; Indels 23; Gaps 21;								
DB	2226 ADEKKQHLIAEKLKERE-REND-SIKDKVENL-ERELQMSSENOELVILDAENSKAEVET	2282							
QY	120 SODGNHNOFTTKVIGKSDRDPDISPKVGENLVGDVVLATGSDOFFEVRSIGSGKGLAA	179							
DB	2283 LKTOIEEVARSLKIFELDLVLRSEKREULTQIOE-KQGHSELDK-LLSSEKSLIEKE	2340							
OY	180 GKY-TDAVTIVTSNOEIRAIRAEONPOFIDLEKVRVEQLTGHGSSVLEELVOLVAKDN	238							
DB	2341 QAEIQIKSESTAVEMLQNLQKLENEAVAAACGDOEIKKATQESLDPIIEEHOLRNSIE	2400							

 M P E S E H
 (TM)

Release 2.1D John F. Collins, BioComputing Research Unit.
 Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
 Distribution rights by IntelliGenetics, Inc.

Mparch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Fri Apr 25 18:54:05 1997; Maspar time 140.57 Seconds
 971.672 Million cell updates/sec

Tabular output not generated.

Title: >US-08-699-716A-1
 Description: (1-1566) from US08699716A.seq
 Perfect Score: 1566
 N.A. Sequence: 1 ATGCGGCATCATCATCA.....ATCACAGCTGTGTAATGA 1566
 Comp: TACCGGTAGTAGTAGTAGT.....TACTGTGCAGACCATTTACT

Scoring table: TABLE default
 Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 113505 seqs, 43611913 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: n-geneseq25
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22

Statistics: Mean 9.447; Variance 5.762; scale 1.639

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	976	62.3	1014	17	T04222	Partial LcrV (V anti
2	972	62.1	1014	17	T04223	Partial LcrV (V anti
3	516	33.0	542	15	092819	Yersinia pestis cafI
4	451	28.8	541	15	092817	Yersinia pestis cafI
5	449	28.7	542	15	092818	Yersinia pestis cafI
6	84	5.4	1047	2	Q10572	Human Natriuretic Pe
7	68	4.3	861	14	081500	sfv anti-rev sequence
8	66	4.2	1047	2	Q10572	Human Natriuretic Pe
9	53	3.4	1402	19	T29246	Human Natriuretic Pe
10	42	2.7	204	1	N81164	Type A neurotoxin C f
11	41	2.6	91	9	Q51746	Base substituted E.co
12	41	2.6	204	1	N81164	Oligonucleotide probe
13	38	2.4	91	9	Q51746	Base substituted E.co
14	37	2.4	114	12	Q70467	Oligonucleotide probe
15	36	2.3	114	12	Q70469	Generic DNA sequence
16	34	2.2	114	12	Q70468	Generic DNA sequence
17	34	2.2	114	12	Q70465	Generic DNA sequence

c	18	33	2.1	114	12	Q70466	Generic DNA sequence	3.90e-05
c	19	32	2.0	114	12	Q70465	Generic DNA sequence	1.30e-04
c	20	32	2.0	114	12	Q70467	Generic DNA sequence	1.30e-04
c	21	31	2.0	114	12	Q70469	Generic DNA sequence	4.28e-04
c	22	31	2.0	114	12	Q70468	Generic DNA sequence	4.28e-04
c	23	30	1.9	36	2	Q11195	Ballast Constituent c	1.39e-03
c	24	29	1.9	39	7	Q51787	Mixed oligonucleotide	4.44e-03
c	25	30	1.9	114	12	Q70470	Generic DNA sequence	1.39e-03
c	26	30	1.9	114	12	Q70465	Generic DNA sequence	1.39e-03
c	27	30	1.9	114	12	Q70470	Generic DNA sequence	1.39e-03
c	28	30	1.9	416	8	Q59427	Human brain Expressed	4.44e-03
c	29	29	1.9	598	1	N90919	Type A Insertion gene	4.44e-03
c	30	29	1.9	1021	16	T04613	5' flanking region of	4.44e-03
c	31	29	1.9	2359	10	Q58609	Sequence encoding Cry	4.44e-03
c	32	29	1.9	4641	1	N81538	Sequence of a gene fr	4.44e-03
c	33	30	1.9	8920	11	Q62924	Carbamoyl-phosphate-s	1.39e-03
c	34	28	1.8	39	18	T05194	Human/mouse guanylate	1.39e-02
c	35	28	1.8	88	18	T17965	Plasmid pET15b His-Ta	1.39e-02
c	36	28	1.8	287	21	T16383	pET-15b expression cl	1.39e-02
c	37	27	1.7	33	2	Q11197	Ballast Constituent c	4.31e-02
c	38	27	1.7	33	7	Q46603	Mixed oligonucleotide	4.31e-02
c	39	26	1.7	42	3	Q14326	MCPC 603 VH CDR2 walk	1.31e-01
c	40	27	1.7	114	12	Q70472	Generic DNA sequence	1.31e-01
c	41	26	1.7	114	12	Q70471	Generic DNA sequence	1.31e-01
c	42	26	1.7	135	22	T25107	Human gene signature	1.31e-01
c	43	27	1.7	250	20	T22591	Human gene signature	4.31e-02
c	44	27	1.7	1150	12	Q70736	TvTA-binding protein-	4.31e-02
c	45	27	1.7	6924	12	Q67283	Human Rmp-beta.	4.31e-02

ALIGNMENTS

RESULT	1	T04222 standard; DNA; 1014 BP.
ID	T04222	18-APR-1996 (first entry)
AC	T04222	Partial lcrV (V antigen) gene of Y. pestis.
DE	Partial lcrV (V antigen) gene of Y. pestis.	
KW	LcrV; V antigen; virulence; plague; vaccine; epitope; ss.	
OS	Yersinia pestis.	
FH	Key	Location/Qualifiers
FT	CDS	1..990
FT	/*tag= a	
FT	/note= "V antigen"	
PN	W09524475-A1.	
PD	14-SEP-1995.	
PF	06-MAR-1995: G00481.	
PR	08-MAR-1994: GB-004577.	
PA	(MINA) UK SEC FOR DEFENCE.	
PI	Leary SEC, Tibball RW, Williamson ED, Leary SE;	
DR	WPI; 95-328268/42.	
DR	P-PSDB: R79961.	
PT	Recombinant DNA expressing Yersinia pestis V antigen - useful in	
PT	oral or parenteral vaccines for protection against plague	
PS	Claim 6, Page 11-13; 25pp: English.	
CC	T04222-23 are DNA sequences (lcrV) encoding all or a protective	
CC	part of the mature V protein of Yersinia pestis. The protein was	
CC	expressed as a fusion protein with maltose binding protein or	
CC	glutathione-S-transferase in 3 different plasmid vectors. Y. pestis is	
CC	the highly virulent causative organism of plague in a wide range of	
CC	animals, including man. The V antigen (lcrV) is an unstable 37.3 kDa	
CC	monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen	
CC	is postulated to act as a virulence antigen, and transformed	
CC	microorganisms contg. recombinant DNA encoding a V antigen protein/	
CC	peptide are useful in vaccines to protect against plague.	
SQ	Sequence 1014 BP; 346 A; 181 C; 201 G; 266 T;	
Query Match	62.3%; Score 976; DB 17; Length 1014;	
Best Local Similarity	99.9%; Pred. No. 0.00e+00;	
Matches	977; Conservative 0; Mismatches 1; Indels 0; Caps 0;	
Db	13 attagagcctacgaacaaacccacacatttatgtgagatctagaanaagttaggctg 72	
QY	589 ATTAGAGCCTACGAAACAAACCCACACATTATTATGAGAGCTGAGAAAAAGTTAGGCTG 648	

D	b	73	gaacaacttactggtcgaatggtcttccagttttagaagaattggtcgaatgtagtcaagaat	132
O	y	649	GAACAACCTTACTGGTCTCATGGTCTTCCAGTTTGTGAAGAATTGGTTCAGTTACTCCAAAGAT	708
D	b	133	aaaataatagatatttcctcatataatagatcccgagaagaattcggaggtttttccaat	192
O	y	709	AAAATATAGATATTTCCTCATTAATATGATCCAGAAAAAGATTCCGAGGTTTTTCCAAAT	768
D	b	193	agagtaattactgtagatcgaattggtcgaagaaatccctagcttattttctacccgag	252
O	y	769	AGAGTAATTACTGATGATATCGAATTGCTCAAGAAATCTCGATTAATTTTCTACCGAG	828
D	b	253	gatgccattctctaaggcggctcatatagaacaaccatctgcaaatctggatcaagagata	312
O	y	829	GATGCCATTCTTAAAGCGGCTCATATGAGAACCAACTCCAAATGGCATCAAGGAGATA	888
D	b	313	aaagagttcccttgatcatcgcgcgaataacaaatggagaattcggcggtctcagtagta	372
O	y	889	AAAGAGTTCCCTTGATCATATGCGCGAATACACATGGGAATTCCGGGCGTTCAATGSCATA	948
D	b	373	atgcaattctcttcaaccgcgcgacgcatcgatcgatgataatttggaaatgattgtagt	432
O	y	949	ATGCAATTCTCTTTAACCGCGCATCGATGATGATGATTTTGGAAATGATTTGTGAT	1008
D	b	433	tcaattgatacatcagtgtagtgcgcgtagaagaatttcgtagaagaattgctgaagttcc	492
O	y	1009	TCAATGATTCATCATGGTGTGATGCCCGAGCAATTCCTGGAAGAAATTACTGATGCTTACC	1068
D	b	493	gcgcaatbaagaatttatacagttatcgaagccgaataatataagcatctgctagtagt	552
O	y	1069	GCCGAAATTAAGATTTATTCAGTTATTCAGGCCGAATTAATAGATCTGCTAGTAGT	1128
D	b	553	ggcaccataaataatccatgataaataccatatactcattcgaataaaattatataggttat	612
O	y	1129	GGCACCAATTAATATCCATGATTAATCCATTAACTCATGATGATAAAATTATATGTGTAT	1188
D	b	613	acagatgaagagatttttaagccagcgacaggtacaaatctcgaagaaatgtgctcaa	672
O	y	1189	ACAGATGAAGAGATTTTAAAGCCAGCGCAGATACAAAATTCTCGAGAAATGCTCTGAA	1248
D	b	673	accaccattcaggtgtagtggagagcagaaaaaatagtctcgataaagacttcttggta	732
O	y	1249	ACCACCAATTCAGTGGATGGAGGAGCAAAAAATATGTCTCGATTAAGAGACTTCTTGGA	1308
D	b	733	agtgagaataaagaaccggggcggttgggtgaatctgaaaaactcatactctataataaa	792
O	y	1309	AGTGAATTAATAAACAACGGGGCGTGGGTAACTGTAACCAAAACCTCAATACCTTATATATAA	1368
D	b	793	gataataatgattatctcactctgtccacacactgtctcgataaagccagcgctcaac	852
O	y	1369	GATAATTAATGAATTAATCTCACTTTGCCACCACTGCTCGATTAAGTCCAGCGCGCTCAAC	1428
D	b	853	gacttggtagccaaaaaacaactcagctggtctgatalatcacacggtttaatcagct	912
O	y	1429	GACTTGGTTAGCCAAAAACAACACACAGCTGTGTGATTAATACACAGTTTATATTCAGCT	1488
D	b	913	attgaagcactcgaaccggtttcattcagaataatagattcagtgatgcgaagcttctgaat	972
O	y	1489	ATTGAAGCAACGAACCGCTTCTATTCAGAAATATGATTAAGTATGCAACGTCCTGCTAAT	1548
D	b	973	gacacgtctgtagtaataa 990	
O	y	1549	GACACGTCCTGTAATATGA 1566	
RESULT 2				
ID	T04223 standard; DNA; 1014 BP.			
AC	T04223:			
DT	18-APR-1996 (first entry)			
DE	Partial LcrV (V antigen) gene of Y. pestis.			
KN	LcrV; V antigen; virulence; plague; vaccine; epitope; ss.			
OS	yersinia pestis.			

Key	Location/Qualifiers
FT CDS	1..990
FT /tag= a	
PN WO9524475-A1.	
PD 14-SEP-1995.	
PF 06-MAR-1995: G00481.	
PA 08-MAR-1994: GB-004577.	
PR (MINA) UK SEC FOR DEFENCE.	
PI Leary SEC, Tithall RW, Williamson ED, Leary SE;	
DR P-PSDB: R79962.	
PT Recombinant DNA expressing Yersinia pestis V antigen - useful in	
PT oral or parenteral vaccines for protection against plague	
PS Cl22.6; Page 15-16; 25pp: English.	
CC T04221-23 are DNA sequences (lcrV) encoding all or a protective epitope	
CC part of the mature V protein of Yersinia pestis. The protein was	
CC expressed as a fusion protein with maltose binding protein or	
CC glutathione-S-transferase in 3 different plasmid vectors. Y. pestis	
CC is the highly virulent causative organism of plague in a wide range of	
CC animals, including man. The V antigen (lcrV) is an unstable 37.3 kDa	
CC monomeric peptide encoded on the ca. 70 kb lcr plasmid. The V antigen	
CC is postulated to act as a virulence antigen, and transformed	
CC microorganisms conty. recombinant DNA encoding a V antigen protein/	
CC peptide are useful in vaccines to protect against plague.	
SQ Sequence 1014 bp; 343 A; 185 C; 204 G; 282 T;	
Query Match 62.1%; Score 972; DB 17; Length 1014;	
Best Local Similarity 99.7%; Pred. No. 0.00e+00;	
Matches 975; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Db 13 atcgaagcctaagacaacaaacacaaatatttattgaagatctcagaanaagttaaggtg 72	
Qy 589 ATTAAAGCCTAAGAACAAACCCACACATTTTATGAGATCTAGAAAAGTTAGGGTG 648	
Db 73 gaacaacttaacgfcgcatggtttcttcagttttagaagaattggttcagttagtaagaat 132	
Qy 649 GAACAACTTACTGGTCAGTCAGTTCTTCACTTTTACAGACATTGGTTCAGTTAGTCAAGAT 708	
Db 133 aaaaaatagatatatttccattaaatgatgatccagaanaagaattcggaggttttgcacat 192	
Qy 709 AAAAATATGATATTTCCATTAAATATGATCCAGAAAAGATCGAGGTTTGGCCAT 768	
Db 193 agagtaattactgatgatatacgaattgtctcaagaanaactcagcttattttccaccgag 252	
Qy 769 AAGATTAATTACTGAGATGATGCAATGCTCAAGAAATCCATGCTTTTCTACCCGAG 828	
Db 253 gatgcattctttaaagcgcgatcatatgatcaacaacatcggaatgatcatcaaacgagta 312	
Qy 829 GATACCATTCTTAAAGGGGCTCATTTATGACAAACCACTGCAAAATGGCATCAACCGAGTA 888	
Db 313 aagaagttccttgaatcatcgccggaatacaaatgggaattgcgcggttcattgcagta 372	
Qy 889 AAGAGTTCCTTGATCATTCGCCGAATACAAATGGGAATTGGCGGCTTCATGCGAGTA 948	
Db 373 atgcatcttctttaaaccgacgatcgtaatcgatgaatatttgaagaatgattgttgat 432	
Qy 949 ATGCATTTCTCTTACCCGCCGATGATGATGATGATTTTGAAGAATGATGTTGAT 1008	
Db 433 tcaatgaatcatcatggtgatggtccggtgacaagtgtggtgaagaattagctcagcttacc 492	
Qy 1009 TCAATGAATCATCATGATGATGCCGTGACAAAGTTGGCTGAAGAAATTAAGTACCTTACC 1068	
Db 493 gccgaattaaagatttactcagttactcgaacgcgaatattaagaactcgtctcagtagt 552	
Qy 1069 GCCGATTTAAGATTATTCAGTTATTCACCGCAATTAATTAAGCATCTGTCTAGTAGT 1128	
Db 553 ggcaccataaatatcatgatataatcaatlaatcatgatgataaaaaattatagtgtat 612	
Qy 1129 GGCACCATTAATATCCATGATATAATCCATTAATCTCATGATATAAAATTTATATGGTTAT 1188	
Db 613 acagatgaagagatttttaaaagcagcgcgagtagtacaatattccggagaatgtctctcaa 672	
Qy 1189 ACAGATGAAGAAATTTTAAAGCCAGCGCCAGATGACAAAAATTCGAGAAATGCCCTCAA 1248	

```

Db 673 accaccatgaagtgtgagcgagagaaaaaatagcttcgataaagagacttcttga 732
    |||||||
QY 1249 ACCACCACTTGAAGTGTGGAGCGAGAAAAAATAGCTTCGATTAAGAGACTTCTTGA 1308
    |||||||
Db 733 agtgaagataaagaacggcggttgggtgaatcgtgaataacatcactctataataa 732
    |||||||
QY 1309 AGTGAGATATAAAAGACCGGGCGTGGTGAATGTGAAAAATCATATCTTTATTAATAA 1368
    |||||||
Db 793 gataataatgaatlatcacttgcacacacactgcgtcgataaagtcaggcgctcaac 852
    |||||||
QY 1369 GATTAATATGAATATATCTCACTTTGCCACACCTGCTGGATTAAGTCAGGCGCTCAAC 1428
    |||||||
Db 853 gaacttggtgcacaaacaaactcagctgtctgatatcattcaatcagtttaactagct 912
    |||||||
QY 1429 GACTTGGTTAGCCAAAAAACACCTGCTGATTAATCATCATCCTTTTAATTCAGCT 1488
    |||||||
Db 913 attgaagcactgaacgcttcattcagaataatgatcattgataagtcagcttctagat 972
    |||||||
QY 1489 ATTGAACCACTGAACCTTTCAATTCAAGAAATATGATTCAAGTGAACGCTGCTAGAT 1548
    |||||||
Db 973 gacacgtctgtaaatga 990
    |||||||
QY 1549 GACACGCTCTGTAAATGA 1566
    |||||||

RESULT 3
ID 092819 standard; DNA; 542 BP.
AC 092819;
DE 17-DEC-1995 (first entry)
DE Yersinia pestis cafI (FI) antigen in plasmid pF0R1b.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
KW bubonic plague; pneumonic plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 2..7
FT /tag= a
FT /note= "first protein encoded by pF0R1b"
FT misc-feature 1..6
FT /tag= b
FT misc-feature 536..541
FT /tag= c
FT /note= "cafI open reading frame downstream seq."
FT CDS 21..530
FT /tag= d
FT /note= "cafI fusion protein"
FT PN W09518231-A1.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA ) UK SEC FOR DEFENCE.
PI Howells A, leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 95-246396/32.
DR P-PSDB: R76528.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Disclosure; Page 19-20; 27pp; English.
CC The sequence represents the plasmid pF0R1b including the entire
CC Y. pestis cafI (FI) antigen gene having a 5' tail including a SacI
CC restriction site, and up to TATAG downstream of the cafI ORF.
CC The DNA construct can be used to transform human or animal gut
CC colonizing microorganisms, specifically attenuated Salmonella
CC typhimurium or Salmonella typhi. The transformed microorganisms
CC can be used as live/attenuated vaccines which induce immune
CC responses at mucosal surfaces. The vaccines provide protection
CC against infection with Y. pestis, and are parenterally and orally
CC active vaccines offering protection against bubonic and pneumonic
CC plague.
SQ Sequence 542 BP; 164 A; 118 C; 114 G; 146 T;

```

Query Match 33.0%; Score 516; DB 15; Length 542;
 Best Local Similarity 99.6%; Pred. No. 0.00e+00;

```

Matches 518: Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 19 atatgaaaaaalcagtlccgltatcgcattgcatlatttggaaactatgcaatgctcia 78
    |||||||
QY 68 ATATGAAAAAATCACTTCGTTATCGCATTCATTTATTTGAACTATGCAACTGTCTA 127
    |||||||
Db 79 atgcggcagaatttaactgcaagaccactgcgaacggcgaactctgttgaacccgccga 138
    |||||||
QY 128 ATGCGGACAGATTTAACTGCAAGCACACCTGCAACGGCAACTCTTTGTGAACGCGCCGA 187
    |||||||
Db 139 tcaacttacataaaggaagcgctccaaattcaatttggcaatggaaatcgata 198
    |||||||
QY 188 TCACCTTTACATATTAAGAGAGGCGCTCCAAATTAATGACATGGAACATGATGATA 247
    |||||||
Db 199 cagaattactgttggtaagccttactcttgcggcctataaaaaaggaacacatgacat 258
    |||||||
QY 248 CAGAAATTACTTGTGTGATAGCTTACTTGGCGGCTATAAACAGACACACTAGACAT 307
    |||||||
Db 259 ctgttaactttacagatgcgcgggtgataccatgtaacttaacttactctagatg 318
    |||||||
QY 308 CTGTTAACTTTACAGATGCGCGCGGTGATCCATGTAATTAACTTACTTCTCAGGATG 367
    |||||||
Db 319 gaaataacacacacactcactacaaaagtatgcaagattcctaagatttgatatc 378
    |||||||
QY 368 GAAATTAACCAACCAATTCACACAAAAGTGAATGCAAGATTCTTAAGATTTGATATCT 427
    |||||||
Db 379 ctccaaagtaaacgggtgagaacctgtgtgggagtgcgtctgttgcaggcagcc 438
    |||||||
QY 428 CTCCTAAGGTAAACGGGTGAACACTTGTGGGGATGACGCTGCTGGCTACGGGACGC 487
    |||||||
Db 439 aggaattcttgttcgcctcaattgtgtccaaagcggtaaacttgcagcaggtaaata 498
    |||||||
QY 488 AGGATTTCTTTGTTCCTCATTCATTTGTTCCAAAGCGGTTAACTTCGACAGGTAATA 547
    |||||||
Db 499 ctgatgctgaaccgtaaccgtaactcacaacaataacatc 538
    |||||||
QY 548 CTGATGCTGTAAACCGTAACCGTATCTTAACCAAGAAATTCAT 587
    |||||||

RESULT 4
ID 092817 standard; DNA; 541 BP.
AC 092817;
DE 17-DEC-1995 (first entry)
DE Yersinia pestis cafI (FI) antigen in plasmid pFGAL2A.
KW Vaccine; antigen; Salmonella typhimurium; Salmonella typhi;
KW bubonic plague; pneumonic plague; ds.
OS Yersinia pestis.
FH Key Location/Qualifiers
FT CDS 2..454
FT /tag= a
FT /note= "vector pFGAL2A bases"
FT PN W09518231-A1.
PF 23-DEC-1994; G02818.
PR 24-DEC-1993; GB-026425.
PA (MINA ) UK SEC FOR DEFENCE.
PI Howells A, leary SEC, Oyston PCF, Titball RW, Williamson ED;
DR WPI: 95-246396/32.
DR P-PSDB: R76526.
PT DNA constructs capable of transforming microorganisms - which can be
PT used as live or attenuated vaccines which induce an immune response,
PT against Yersinia pestis, at mucosal surfaces.
PS Claim 7; Page 15-16; 27pp; English.
CC The sequence represents the plasmid pFGAL2A construct showing the
CC fusion of the first few bases of beta-galactosidase in the vector
CC with the Y. pestis cafI (FI) antigen minus its signal sequence and
CC having a 5' tail including a SacI restriction site, and up to the
CC cafI AAC-3' end with some vector bases. The DNA construct can be

```

CC used to transform human or animal gut colonizing microorganisms,
CC specifically attenuated *Salmonella typhimurium* or *Salmonella typhi*
CC The transformed microorganisms can be used as live/attenuated
CC vaccines which induce immune responses at mucosal surfaces. The
CC vaccines provide protection against infection with *Y. pestis*, and
CC are parenterally and orally active vaccines offering protection
CC against bubonic and pneumonic plague.
50 Sequence 541 BP; 163 A; 120 C; 111 G; 147 T;

Query Match	28.8;	Score 451;	DB 15;	Length 541;
Best Local Similarity	99.3%;	Pred. No. 1.99e-286;		
Matches 454; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Db	6	CGCAGATTACGTCGAAGCAGCCACTGCAAGGGCAACTCTGTTGAACGAGCCCGCATCA	65
Oy	131	CGGCAATTATTACTGCAAGCACCCACTGCAGAGGCAACTCTGTGGACCAAGCCGCATCA	190
Db	66	CTATCACATATAAGGAAGCGCTCCCAATCTACAATCTGACAAATGGAACATCGATAG	125
Oy	191	CTCTTACATATATAGGAAGGGCCCTCCATTACAAATTATGGACATATGGAAATCATGATATCAAG	250
Db	126	AATCACTGTGTGTAAGCTACTCTCTTGGCGGTATAAACAAGAAACCAATAGCAATCTG	185
Oy	251	AATTCTGTGTGGATACGGTTACTCTTTGGCGGCTATTAACAGGACCACTACGACATCTAG	310
Db	186	TTAACTTACAGATGCGCGGGATGATCCATAGTAACTTAACATTAATCTTCACAGATGGAA	245
Oy	311	TTTAACTTACAGATGATCCCGGGTATGCCATGATCTTAAATTAATTAATCTTCACAGATGGAA	370
Db	246	ATAACCAACCAATCTACTACAAGAATGTATGGCAAGATCTTAGAGATTTGATATCTCTC	305
Oy	371	ATAACCAACCAATTTCACTACCAAAATGATTTGGCAAGATCTTAGAGATTTGATATCTCTC	430
Db	306	CTAAGGTAAGAAGTGGAAAGCTGTGGGGGATGACGTCGCTTGGCTACAGGCAAGCCAG	365
Oy	431	CTAAGGTAAGAAGTGGAAAGCTTGTGGGGGATGACGTCGCTTGGCTACAGGCAAGCCAG	490
Db	366	ATTCTTGTCTGCATCAATGTGTCACAAGCGGTAACTGCAAGCAAGTAAATACACTG	425
Oy	491	ATTCTTGTCTGCATCAATGTGTCCAAAGGGCGTAACTTGGACGACGATTAATACACTG	550
Db	426	ATGCTGTAAACCGTAAACCGTATCTACCAATATCAT	462
Oy	551	ATGCTGTAAACCGTAAACCGTATCTAAACAGAAATTCAT	587

```

RESULT      5
ID          092818 standard; DNA: 542 BP.
AC          Q92818;
DT          17-DEC-1995 (first entry)
DE          Yersinia pestis catII (Pl) antigen in plasmid pF513a.
KW          Vaccinia; antigen; Salmonella typhimurium; Salmonella typhi;
           bubonic plague; pneumonic plague; ds.
OS          Yersinia pestis.
FH          Key
FT          CDS          Location/Qualifiers
FT          /*tag=       a
FT          misc_feature 1..6
FT          /*tag=       b
FT          /note= "E. coli LTB signal peptide bases"
FT          misc_feature 536..541
FT          /*tag=       c
FT          /note= "vector pF513a bases"
PN          W09518231-A1.
PD          06-JUL-1995.
PE          23-DEC-1994; G02818.
PR          24-DEC-1993; GB-026425.
PA          (MUNA ) UK SEC FOR DEFENCE.
PI          Howells A. Leary SEC, Oyston PCR, Tibball RW, Williamson ED;
           Wpl: 95-246396/32.
P           P-FSD8: R76527.
DR          DNA constructs capable of transforming microorganisms - which can be
           used as live or attenuated vaccines which induce an immune response,
           PT

```

PT against *Yersinia pestis*, at mucosal surfaces.
PS Disclosure; Page 17-18; 27pp; English.

CC The sequence represents the plasmid pPSIG3a construct showing the
CC fusion of the first few bases of the *E. coli* LTR signal sequence
CC with the *Y. pestis* catI (FI) antigen minus its signal sequence and
CC having a 5' tail including a SacI restriction site, and up to the
CC catI AAC-3' end with some vector bases. The DNA construct can be
CC used to transform human or animal gut colonizing microorganisms,
CC specifically attenuated *Salmonella typhimurium* or *Salmonella typhi*
CC The transformed microorganisms can be used as live/attenuated
CC vaccines which induce immune responses at mucosal surfaces. The
CC vaccines provide protection against infection with *Y. pestis*, and
CC are parenterally and orally active vaccines offering protection
CC against bubonic and pneumonic plague.
50 Sequence 542 BP; 163 A; 122 C; 110 G; 147 T;

Query Match	28.7%;	Score 449;	DB 15;	Length 542;
Best Local Similarity	99.3%;	Pred. No. 4,94e-285;		
Matches 452; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

Db	9	gcagatthaactggaagcagccacgcggaaggaactctttgaacagccgcacacac	68
Oy	133	GCAGATTTAACTGGAAGACGACCTACGCAACGGCAACTCTTTGTAACCGCCGCACATCACT	192
Db	69	attacataaaggaagcgccgcccaattacaattaigacaatlgaaacatcgatacaga	128
Oy	193	CTTACATTTAAGGAAGGCGCTCCCAATTACATTTATGACACATGGAACATCGATTACAGA	252
Db	129	ttaactgtgtgtagcgtctactctctggcggctataaagaagcaaccagtagacatcgtt	188
Oy	253	TTACTGTGTGGTAGCGCTTACTCTTGGGGGCTATTTAAACAGAAACCATACACATCTGTT	312
Db	189	aacttiacaagtcgcggcggtgatacccatctactiaacatctiaacttcagatgggaat	248
Oy	313	AACTTTACAGATGCCCGGGGATGCCCATGTACTTAACTTACTCTCGAGATGCAAT	372
Db	249	aaccaccaatctacatacaaaagtattgycgaagatctagagatttgatalctctct	308
Oy	373	AACCAACCATTCACATACAAAAGTATGTCGAAGATCTAGAAATTTGATATCTCTCT	432
Db	309	aagttaaacggtgagaacctgtgtggagatgaactgctctgtctcagggcagaacagat	368
Oy	433	AAGGTAACGGGTGGAACCTGTGTGGGAATGACCTCCTCTTGGCTACGGGCACCCAGAT	492
Db	369	tctctgtgtcgtcaatctgtgtccaaagcggtaaactctcagcaggtlaataacatgat	428
Oy	493	TTCTTTGTTGCCTCAATTTGGTTCAAAGCGGGTAACTTCAGAGGTAATACATGAT	552
Db	429	gctgtaaacgtaaacccgatctcaacacataatccat	463
Oy	553	GCTGTAAACCGTAACCGTATCTTAAACAAGAAATTCAT	587

ID	RESULT	6
AC	010572: standard; DNA; 1047 Bp.	
DT	09-APR-1991 (first entry)	
DE	Human Natriuretic Peptide Receptor B.	
NPB:	ANP; BNP; CNP; kidney failure; heart failure; protein kinase.	
KW	hyperaldosteronism; glaucoma; guanyl cyclase.	
OS	Homo sapiens.	
FH	Key	location/Qualifiers
FT	Peptide	1..22
FT	/label=	signal sequence
FT	protein	12
FT	/label=	mature NPB
FT	Domain	23..45
FT	/label=	extracellular domain
FT	/note=	"binds natriuretic peptides A,B and C"
FT	Domain	456..456
FT	/label=	transmembrane domain
FT	Domain	479..1047
FT	/label=	cytoplasmic domain


```

FI  comments) "
PN  WO9418318-A.
PD  18-AUG-1994.
PF  01-FEB-1994; U00977.
PR  01-FEB-1993; US-013416.
PR  30-DEC-1993; US-176500.
PR  31-JAN-1994; US-189331.
PA  (DYNC-) UNIT NORTH CAROLINA.
PI  Fowlkes DW, Kay BK;
DR  WPI; 94-279739/34.
PT  Identifying proteins or peptide(s) which bind a ligand - by
PT  screening a recombinant vector library expressing fusion proteins
PT  comprising a binding domain and an effector domain
PS  Disclosure; Page 35; 255pp; English.
CC  Q70469 is a generic DNA sequence used to generate random TSAR peptide
CC  This generic formula can be represented as follows: X(TGC)(NNB)10-
CC  (TGC)(NNB)6Z(NNB)2(TGC)(NNB)14(TGC)Y. X and Y are flanking restriction
CC  sites (X is not the same as Y) that are not specified further. This
CC  sequence generates peptides that are cloverleaf in structure. Other
CC  generic sequences are shown in Q70465-68. Other specific peptides
CC  generated by these generic sequences are shown in R65150-54. TSARs are
CC  concatenated heterofunctional proteins or peptides, comprising at least
CC  two functional regions - a binding domain with affinity for a ligand and
CC  a second effector peptide portion that is chemically or biologically
CC  active. They may further comprise a linker peptide between the 2 domains.
CC  The oligonucleotides are also designed so that the expressed peptide
CC  contains 2 or 4 cysteine residues positioned in, or flanking, the
CC  unpredicted or variant residues. These residues confer some degree of
CC  conformational rigidity to the peptides. The TSARs or compsns. comprising
CC  a TSAR binding domain can be used in vivo to deliver a chemically or
CC  biologically active moiety, eg. metal ion, radioisotope, peptide, toxin
CC  or enzyme, to the specific target or on the cell. They can also replace
CC  the function of macromolecules, eg. monoclonal or polyclonal antibodies
CC  and therefore circumvent the need for complex methods of hybridoma
CC  formation or in vivo antibody production. The TSARs are easily
CC  characterised and have designed activity allowing direct and rapid
CC  detection in a screening process.
SQ  Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

Query Match 2.3%; Score 36; DB 12; Length 114;
Best Local Similarity 8.0%; Pred. No. 9.77e-07;
Matches 9; Conservative 30; Mismatches 73; Indels 0; Gaps 0;

Db 3 cdbnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnbhnnnnnn 62
   | : : : : : : : : : : : : : : : : : : : : : : : :
CP 229 CCATATTGTAATGGAGCGCCTCTATATAGTAGTGAAGCGCGCTGTTCAACAA 170
   : : : : : : : : : : : : : : : : : : : : : : : :

Db 63 bnbhbgcnbhbhbhbhbhbhbhbhbhbhbhbhbhbhbhbhbhbhbhbhb 114
   : : : : : : : : : : : : : : : : : : : : : : : :
CP 169 GAGTGGCCGTGCAAGTGGTCTTGCACTTAATCTGCCGATTAGCAGTTGC 118

```

Search completed: Fri Apr 25 18:56:32 1997
 Job time : 147 secs.

THIS PAGE BLANK (USPTO)